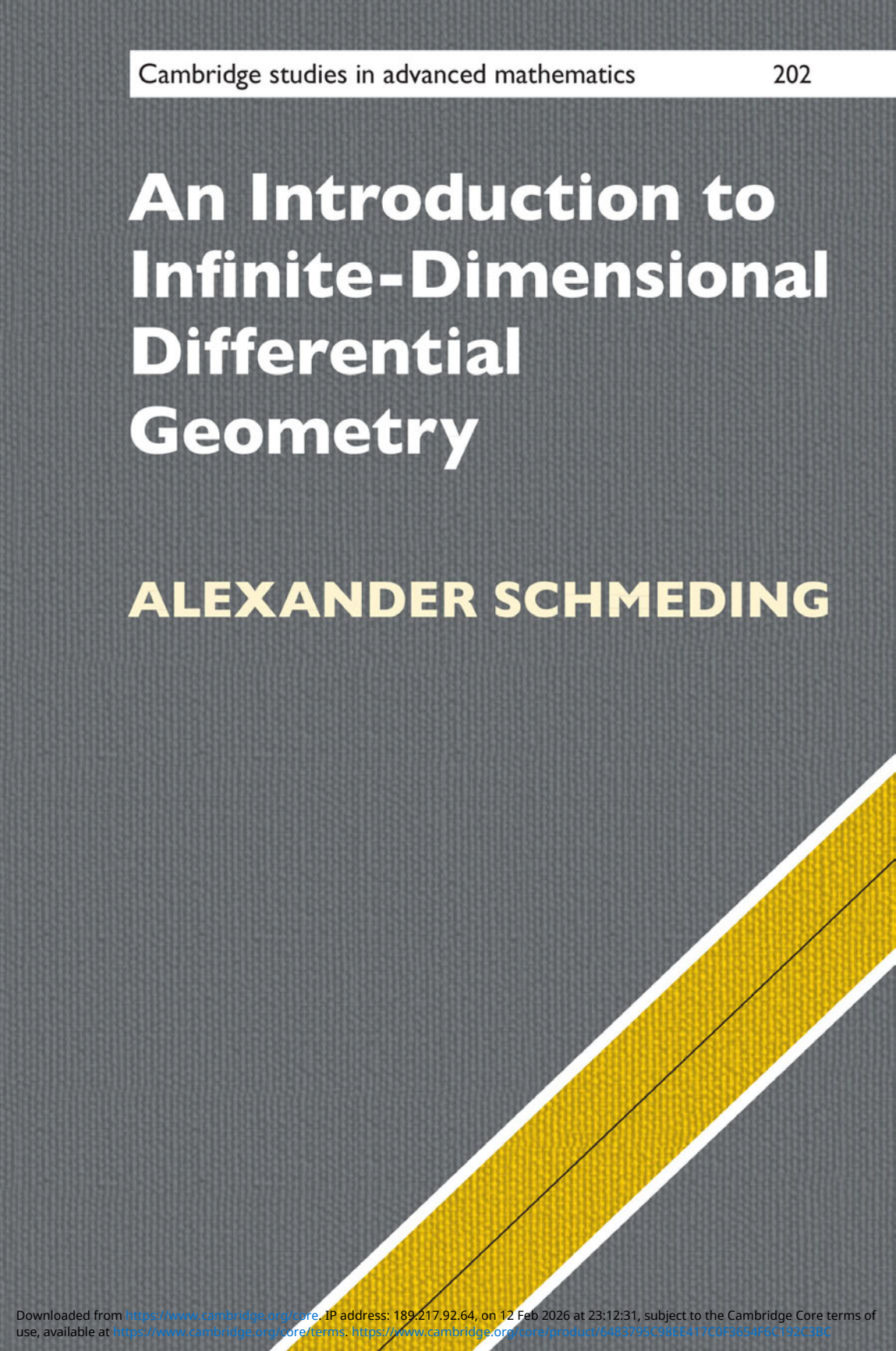
Cambridge studies in advanced mathematics 202

An Introduction to Infinite-Dimensional Differential Geometry

ALEXANDER SCHMEDING



Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6485795C5B6E417C0F3654F6C192C38C

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

CAMBRIDGE STUDIES IN ADVANCED MATHEMATICS 202

Editorial Board
J. BERTOIN, B. BOLLOBÁS, W. FULTON, B. KRA, I. MOERDIJK,
C. PRAEGER, P. SARNAK, B. SIMON, B. TOTARO

AN INTRODUCTION TO INFINITE-DIMENSIONAL DIFFERENTIAL GEOMETRY

Introducing foundational concepts in infinite-dimensional differential geometry beyond Banach manifolds, this text is based on Bastiani calculus. It focuses on two main areas of infinite-dimensional geometry: infinite-dimensional Lie groups and weak Riemannian geometry, exploring their connections to manifolds of (smooth) mappings. Topics covered include diffeomorphism groups, loop groups and Riemannian metrics for shape analysis. Numerous examples highlight both surprising connections between finite- and infinite-dimensional geometry, and challenges occurring solely in infinite dimensions. The geometric techniques developed are then showcased in modern applications of geometry such as geometric hydrodynamics, higher geometry in the guise of Lie groupoids and rough path theory. With plentiful exercises, some with solutions, and worked examples, this will be indispensable for graduate students and researchers working at the intersection of functional analysis, non-linear differential equations and differential geometry. This title is also available as open access on Cambridge Core.

Alexander Schmeding is Associate Professor in Mathematics at Nord University at Levanger.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

CAMBRIDGE STUDIES IN ADVANCED MATHEMATICS 202

Editorial Board
J. Bertoin, B. Bollobás, W. Fulton, B. Kra, I. Moerdijk, C. Praeger, P. Sarnak, B. Simon, B. Totaro

All the titles listed below can be obtained from good booksellers or from Cambridge University Press. For a complete series listing, visit www.cambridge.org/mathematics.

Already Published
162. C. J. Bishop & Y. Peres *Fractals in Probability and Analysis*
163. A. Bovier *Gaussian Processes on Trees*
164. P. Schneider *Galois Representations and* (φ, Γ)*-Modules*
165. P. Gille & T. Szamuely *Central Simple Algebras and Galois Cohomology* (*2nd Edition*)
166. D. Li & H. Queffelec *Introduction to Banach Spaces, I*
167. D. Li & H. Queffelec *Introduction to Banach Spaces, II*
168. J. Carlson, S. Müller-Stach & C. Peters *Period Mappings and Period Domains* (*2nd Edition*)
169. J. M. Landsberg *Geometry and Complexity Theory*
170. J. S. Milne *Algebraic Groups*
171. J. Gough & J. Kupsch *Quantum Fields and Processes*
172. T. Ceccherini-Silberstein, F. Scarabotti & F. Tolli *Discrete Harmonic Analysis*
173. P. Garrett *Modern Analysis of Automorphic Forms by Example, I*
174. P. Garrett *Modern Analysis of Automorphic Forms by Example, II*
175. G. Navarro *Character Theory and the McKay Conjecture*
176. P. Fleig, H. P. A. Gustafsson, A. Kleinschmidt & D. Persson *Eisenstein Series and Automorphic Representations*
177. E. Peterson *Formal Geometry and Bordism Operators*
178. A. Ogus *Lectures on Logarithmic Algebraic Geometry*
179. N. Nikolski *Hardy Spaces*
180. D.-C. Cisinski *Higher Categories and Homotopical Algebra*
181. A. Agrachev, D. Barilari & U. Boscain *A Comprehensive Introduction to Sub-Riemannian Geometry*
182. N. Nikolski *Toeplitz Matrices and Operators*
183. A. Yekutieli *Derived Categories*
184. C. Demeter *Fourier Restriction, Decoupling and Applications*
185. D. Barnes & C. Roitzheim *Foundations of Stable Homotopy Theory*
186. V. Vasyunin & A. Volberg *The Bellman Function Technique in Harmonic Analysis*
187. M. Geck & G. Malle *The Character Theory of Finite Groups of Lie Type*
188. B. Richter *Category Theory for Homotopy Theory*
189. R. Willett & G. Yu *Higher Index Theory*
190. A. Bobrowski *Generators of Markov Chains*
191. D. Cao, S. Peng & S. Yan *Singularly Perturbed Methods for Nonlinear Elliptic Problems*
192. E. Kowalski *An Introduction to Probabilistic Number Theory*
193. V. Gorin *Lectures on Random Lozenge Tilings*
194. E. Riehl & D. Verity *Elements of* ∞*-Category Theory*
195. H. Krause *Homological Theory of Representations*
196. F. Durand & D. Perrin *Dimension Groups and Dynamical Systems*
197. A. Sheffer *Polynomial Methods and Incidence Theory*
198. T. Dobson, A. Malnič & D. Marušič *Symmetry in Graphs*
199. K. S. Kedlaya *p-adic Differential Equations*
200. R. L. Frank, A. Laptev & T. Weidl *Schrödinger Operators:Eigenvalues and Lieb–Thirring Inequalities*
201. J. van Neerven *Functional Analysis*

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

An Introduction to Infinite-Dimensional Differential Geometry

ALEXANDER SCHMEDING
Nord University, Levanger

CAMBRIDGE
UNIVERSITY PRESS

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781316514887
DOI: 10.1017/9781009091251

© Alexander Schmeding 2023

This work is in copyright. It is subject to statutory exceptions and to the provisions of relevant licensing agreements; with the exception of the Creative Commons version the link for which is provided below, no reproduction of any part of this work may take place without the written permission of Cambridge University Press.

An online version of this work is published at doi.org/10.1017/9781009091251 under a Creative Commons Open Access license CC-BY-NC-ND 4.0, which permits re-use, distribution and reproduction in any medium for non-commercial purposes providing appropriate credit to the original work is given, through funding from Nord Universitet, Postboks 1490, 8049 Bodø, Norway. You may not distribute derivative works without permission. To view a copy of this license, visit https://creativecommons.org/licenses/by-nc-nd/4.0

All versions of this work may contain content reproduced under license from third parties. Permission to reproduce this third-party content must be obtained from these third-parties directly.

When citing this work, please include a reference to the DOI 10.1017/9781009091251

First published 2023

A catalogue record for this publication is available from the British Library.

ISBN 978-1-316-51488-7 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Contents

	Preface		*page* vii
1	**Calculus in Locally Convex Spaces**		1
	1.1	Introduction	1
	1.2	Curves in Locally Convex Spaces	3
	1.3	Bastiani Calculus	7
	1.4	Bastiani versus Fréchet Calculus on Banach Spaces	13
	1.5	Infinite-Dimensional Manifolds	16
	1.6	Tangent Spaces and the Tangent Bundle	19
	1.7	Elements of Differential Geometry: Submersions and Immersions	23
2	**Spaces and Manifolds of Smooth Maps**		30
	2.1	Topological Structure of Spaces of Differentiable Mappings	30
	2.2	The Exponential Law and Its Consequences	37
	2.3	Manifolds of Mappings	41
3	**Lifting Geometry to Mapping Spaces I: Lie Groups**		48
	3.1	(Infinite-Dimensional) Lie Groups	48
	3.2	The Lie Algebra of a Lie Group	55
	3.3	Regular Lie Groups and the Exponential Map	62
	3.4	The Current Groups	69
4	**Lifting Geometry to Mapping Spaces II: (Weak) Riemannian Metrics**		80
	4.1	Weak and Strong Riemannian Metrics	80
	4.2	The Geodesic Distance on a Riemannian Manifold	84
	4.3	Geodesics, Sprays and Covariant Derivatives	90
	4.4	Geodesic Completeness and the Hopf–Rinow Theorem	102

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

5 Weak Riemannian Metrics with Applications in Shape Analysis 106
5.1 The L^2-metric and Its Cousins 107
5.2 Shape Analysis via the Square Root Velocity Transform 114

6 Connecting Finite-Dimensional, Infinite-Dimensional and Higher Geometry 120
6.1 Diffeomorphism Groups Determine Their Manifolds 120
6.2 Lie Groupoids and Their Bisections 123
6.3 (Re-)construction of a Lie Groupoid from Its Bisections 133

7 Euler–Arnold Theory: PDEs via Geometry 138
7.1 Introduction 138
7.2 The Euler Equation for an Ideal Fluid 143
7.3 Euler–Poincaré Equations on a Lie Group 146
7.4 An Outlook on Euler–Arnold Theory 150

8 The Geometry of Rough Paths 157
8.1 Introduction 157
8.2 Iterated Integrals and the Tensor Algebra 159
8.3 A Rough Introduction to Rough Paths 168
8.4 Rough Paths and the Shuffle Algebra 176
8.5 The Grand Geometric Picture (Rough Paths and Beyond) 181

Appendix A **A Primer on Topological Vector Spaces and Locally Convex Spaces** 186
Appendix B **Basic Ideas from Topology** 206
Appendix C **Canonical Manifold of Mappings** 213
Appendix D **Vector Fields and Their Lie Bracket** 225
Appendix E **Differential Forms on Infinite-Dimensional Manifolds** 231
Appendix F **Solutions to Selected Exercises** 244

References 256
Index 264

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Preface

Analysis and geometry on infinite-dimensional spaces is an active research field with many applications in mathematics and physics. Examples for applications arise naturally even when one is interested in problems that on first sight seem genuinely finite dimensional. You might have heard that it is impossible to accurately predict the weather over a long time. It turns out that this can be explained by studying the curvature of certain infinite-dimensional manifolds (Arnold, 1966). This example shows that everyday phenomena are intricately linked to geometric objects residing on infinite-dimensional manifolds. In recent years the list of novel applications for infinite-dimensional (differential) geometry has broadened considerably. Among the more surprising novelties are applications in stochastic and rough analysis (rough path theory à la T. Lyons leads to spaces of paths in infinite-dimensional groups; see Friz and Hairer, 2020) and renormalisation of stochastic partial differential equations via Hairer's regularity structures (see Bogfjellmo and Schmeding, 2018).

The aim of this book is to give an introduction to infinite-dimensional (differential) geometry. Differential geometry in infinite dimensions comes in many flavours, such as Riemannian and symplectic geometry. One can study Lie groups and their actions as well as Kähler manifolds, or Finsler geometry. As should already be apparent from this very incomplete list, it is simply not possible to cover a sizeable portion of the diverse topics subsumed under the label (infinite-dimensional) differential geometry. Hence the present book will focus on two main areas: Riemannian geometry and Lie groups. These topics are arguably the most prominent and well studied of the above list. Moreover, certain basic but important examples in both topics can be approached based on basic results on manifolds of (differentiable) mappings. However, it is important to stress that the focus of this book is introductory in nature. We will usually refrain from discussing results in their most general form if this allows

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

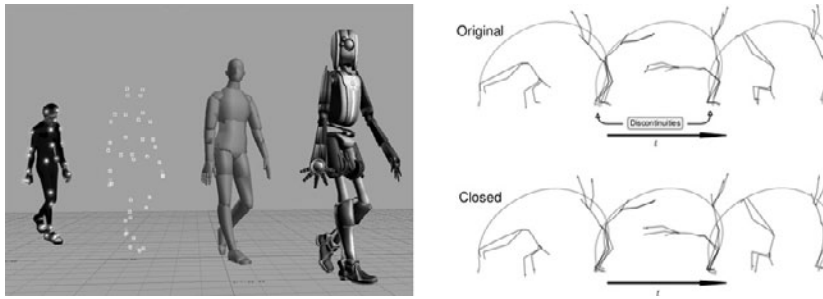


Figure 1 (a) Motion capturing using active markers to animate a figure. Based on the Wikimedia commons picture Activemarker2, public domain; see https://commons.wikimedia.org/wiki/File:Activemarker2.PNG, accessed: 07.12.21. (b) Removing discontinuities from animations via geometric methods. Reproduced from Celledoni et al. (2016) with permission of AIMS.

us to avoid lengthy technical discussions. Before we outline the programme of the book further, let us highlight two applications of infinite-dimensional geometric structures.

Shape analysis. Shapes are unparametrised curves in a vector space/on a manifold. Mathematically these are modelled by considering spaces of differentiable functions and quotienting out an appropriate action of a diffeomorphism group (modelling the reparametrisation). Spaces of differentiable functions can only be modelled using infinitely many parameters, whence they are prime examples of infinite-dimensional spaces. Notice that when talking about spaces of differentiable functions, the functions themselves are thought of as points in the infinite-dimensional space. This is a subtle point as a path between two points (aka functions) in this space will be a curve of curves. The aim of shape analysis is now to compare and transform shapes. For the comparison task one is interested in computing shortest distances between shapes. In the language of Riemannian geometry, one thus wants to compute geodesics (curves that are locally of shortest length) in these spaces. From the geometric formulation, a plethora of applications ranging from medical imaging to computer graphics can be dealt with. As a sample application we would like to highlight the processing of motion capturing data.

The data from the motion capturing process leads to a skeletal wireframe. Since motion capturing is, in general, expensive, one would like to create algorithms that interpolate between different movements or remove discontinuities occuring in the looping of motions (see Figure 1 for a graphical example of the problem and its (numerical) solution).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Current groups and diffeomorphism groups. Groups arising from problems related to differential geometry, fluid dynamics and the symmetry of evolution equations are often naturally infinite-dimensional manifolds with smooth group operations. Prime examples are the diffeomorphism groups $\mathrm{Diff}(K)$ for K a smooth and compact manifold. We encountered them already as reparametrisation groups in the shape analysis example. However, they are also of independent interest, for example in fluid dynamics: if K is a three-dimensional torus, the motion of a particle in a fluid corresponds, under periodic boundary conditions, to a curve in $\mathrm{Diff}(K)$; see Ebin and Marsden (1970). This observation led Arnold to his discovery of a general method by which (certain) partial differential equations (PDEs) can be lifted to ordinary differential equations on diffeomorphism groups. Typically these PDEs arise in the context of hydrodynamics. Many of the applications of this technique, nowadays known as *Euler–Arnold theory for PDEs*, come from geometric hydrodynamics (see Khesin and Wendt, 2009 or Modin, 2019 for an introduction). The already mentioned relation of weather forecasts to infinite-dimensional manifolds arises in a similar fashion.

To deal with these examples, one frequently needs to leave the theory of differential geometry on Banach manifolds (Lang, 1999), since diffeomorphism groups cannot be modelled as Lie groups on Banach spaces. To understand this, consider the canonical action of the diffeomorphisms $\mathrm{Diff}(K)$ on a connected compact manifold K, that is,

$$\alpha\colon\ \mathrm{Diff}(K)\times K\to K,\quad (\varphi,k)\mapsto\varphi(k).$$

This action is effective (i.e. if $\varphi(k)=\psi(k)$ for all k, then $\varphi=\psi$), and with some work one can show that the action α is transitive (i.e. for every pair of points there is a diffeomorphism mapping the first to the second; see Michor and Vizman, 1994). We would expect that any canonical Lie group structure turns α into a smooth map. However, Omori established the following:

Theorem (Omori, 1978) *If a (connected) Banach–Lie group G acts effectively, transitively and smoothly on a compact manifold, then G must be a finite-dimensional Lie group.*

Thus, since there is no way for the diffeomorphism group being described only by finitely many parameters, the diffeomorphism group $\mathrm{Diff}(K)$ cannot be a Lie group. To treat these examples, one thus has to look beyond the realm of Banach spaces.

Another example are the so-called current groups $C^\infty(K,G)$ of smooth mappings from a compact manifold to a Lie group G. The group structure is

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

here given by pointwise multiplication of functions. In physics, current groups describe symmetries in Yang–Mills theories. For example, in the theory of electrodynamics, the gauge transformations for Maxwell's equations form a (Lie) subgroup of a current group. Note that the manifold K in physically relevant theories typically models spacetime and is non-compact. However, we will restrict ourselves in this book to current groups (and function spaces) on compact manifolds K. This allows us to avoid an overly technical discussion while special properties of the main examples remain accessible. As an example, we mention the 'lifting' of geometric features from finite-dimensional target Lie groups to the infinite-dimensional current groups. This procedure works particularly well in the special case where $K = \mathbb{S}^1$ is the unit circle. In this case the current group is better known under the name *loop group* $LG := C^\infty(\mathbb{S}^1, G)$; see Pressley and Segal (1986). Loop groups and current groups are great examples for a general property of spaces of smooth functions which we shall encounter often: under suitable assumptions, the infinite-dimensional geometry arises from 'lifting the finite-dimensional geometry' of the target spaces.

A common topic of the applications and mathematical topics mentioned above is an intimate connection between finite- and infinite-dimensional geometry. Indeed while infinite-dimensional differential geometry might seem like an arcane topic from the perspective of the finite-dimensional geometer, there are many sometimes surprising connections between the realms of finite- and infinite-dimensional geometry. We already mentioned Euler–Arnold theory and Arnold's insights relating curvature on infinite-dimensional manifolds to weather forecasts. Another example is Duistermaat and Kolk's proof of the third Lie theorem ('*every finite-dimensional Lie algebra is the Lie algebra of a Lie group*'; Duistermaat and Kolk, 2000, Section 1.14) or Klingenberg's investigation of closed geodesics (see Klingenberg, 1995). In both cases, infinite-dimensional techniques on path spaces were leveraged to solve the finite-dimensional problems. While many of the examples just mentioned can be studied while staying in the realm of finite-dimensional geometry, the link to infinite-dimensional geometry should not be disregarded as a purely academic exercise. A case in point might be the theory of rough paths discussed in §8. Here it generally suffices to consider truncated and thus finite-dimensional geometric settings. However, the infinite-dimensional limiting objects hint at deeper geometric insights hidden in the finite-dimensional perspective. It is my view that the infinite-dimensional perspective not only provides a convenient framework for these examples but exhibits important underlying structures and principles. These are worth exploring both for their own sake and for the connected applications.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

In the context of the present book, we will explore the connection between the finite- and infinite-dimensional realm in the context of manifolds of differentiable mappings. These manifolds allow the lifting of geometric structures such as Lie groups and Riemannian metrics to interesting infinite-dimensional structures. For finite orders of differentiability, manifolds of differentiable mappings can be modelled on Banach spaces (see Palais, 1968). Thus they are within the grasp of differential geometry on Banach spaces (Lang, 1999). However, since certain constructions (such as the exponential law) become much more technical for finite orders of differentiability, we shall in the present book study exclusively manifolds and spaces of smooth (i.e. infinitely often differentiable) mappings. This places us immediately outside of the realm of Banach manifolds but enables important constructions such as the Lie group structure for diffeomorphism groups.

Finally, I would like to draw the reader's attention to the fact that besides many links and connections between finite- and infinite-dimensional geometry, there are quite severe differences between both settings of differential geometry. Many of the strong structural statements from finite-dimensional geometry simply cease to be valid when passing to manifolds modelled on infinite-dimensional spaces (many statements already break down in the more familiar Banach setting). These pitfalls are well known to experts but often surprise beginners in the field. In the spirit of providing an introduction to infinite-dimensional geometry, I have taken care to emphasise the differences and illustrate them with examples where possible. For example, the following problems are frequently encountered when passing to the infinite-dimensional setting:

- Infinite-dimensional spaces *cannot be locally compact* and they do not support a unique vector topology. Thus arguments building on compactness are not available (see Appendix A).
- Smooth bump functions do not need to exist (see Appendix A.4). Hence the usual local-to-global arguments become unavailable.
- Beyond Banach spaces, there is *no general solution theory for ordinary differential equations* and no general inverse function theorem (see Appendix A.5).
- Dual spaces become more difficult to handle. As a consequence, it is impossible to define differential forms as sections in dual bundles in general (see Remark 1.45).
- Equivalent definitions from finite-dimensional differential geometry are often not equivalent in the infinite-dimensional setting. This happens, for example, for submersions and immersions; compare §1.7.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

The structure of the book is as follows. First of all, we shall provide the necessary foundational material for differential geometry in the general infinite-dimensional setting in the first two chapters. Here we emphasise

(a) calculus and manifolds on infinite-dimensional spaces beyond the Banach setting, and
(b) manifolds of differentiable functions.

Beyond Banach spaces there are several choices as to how one can generalise the concept of differentiability. We adopted the so-called Bastiani calculus based on iterated directional derivatives. This choice is different from the popular 'convenient calculus'. We shall compare the calculi in later chapters. Furthermore, we discuss several foundational topics, such as locally convex spaces, in a series of appendices. The material covered there should allow the reader to grasp most of the basics needed to follow the main part of the book. Moreover, the material in the appendices gives some insight into the common problems arising in the passage from finite- to infinite-dimensional settings already mentioned. While the setting in which we will be working is quite general, we will often not exhibit the most general definitions, results and settings which can possibly be treated in the framework. For example, only manifolds of mappings on a compact manifold are considered here. The non-compact case is interesting and deserving of attention, however the associated theory is much more involved and technical. Our philosophy here is that if the simplified case already admits applications and exhibits the character and problems of the infinite-dimensional setting, we will restrict our attention to the simple case. However, we shall comment on the more general case and provide pointers to the literature. Armed with the knowledge provided by the present book, the reader should be able to quickly learn the more general case should she so desire.

Having dealt with the general setting, we shall then study the main objects of interest in this book:

(a) (infinite-dimensional) Lie groups and Lie algebras, Chapter 3 and
(b) (weak) Riemannian geometry, Chapter 4.

Based on these building blocks, we shall explore several applications of infinite-dimensional geometry in the next chapters. These range from shape analysis to connections with higher geometry (in the guise of Lie groupoids) to Euler–Arnold theory for PDEs and the geometry of rough path spaces. These later chapters can be read mostly independently from each other. I have selected the topics for the advanced chapters with a view towards developments within the field over the last years. The broad selection of topics and the introductory

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

nature of the present book prevent an in-depth discussion of these advanced topics. Therefore these chapters gracefully omit many of the more technical and subtle points. However, the reader will be able to gain an impression of the role infinite-dimensional differential geometry plays in these applications. Moreover, there are ample references to the literature which should enable the interested reader to follow up on a topic after perusing the respective chapter.

Before we begin, let us set some conventions that will be in effect for all that is to come.

Conventions

- Write $\mathbb{N} := \{1, 2, 3, \ldots\}$ for the natural numbers and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.
- All topological spaces are assumed to be Hausdorff, that is, for every pair of two (distinct) points in the space, there are disjoint open neighbourhoods of these points.
- If nothing else is said, products of topological spaces will always carry the product topology.
- If $U \subseteq X$ is an open subset of a topological space, we also write shorter $U \Subset X$.

We shall exclusively work over the real numbers $\mathbb{R}$, all vector spaces are to be understood over the real numbers. However, many results carry over to complex vector spaces. For example, it is no problem to define the notion of complex differentiability (see e.g. Glöckner, 2002).

Recommended Further Reading

As mentioned above, infinite-dimensional differential geometry is a vast topic, which we cannot hope to cover in this book. Thus we have concentrated on an introduction to infinite-dimensional Lie groups, (weak) Riemannian metrics and their interplay. For further reading on the topics of this book, we would like to recommend the following works, which have also influenced the presentation in the present book:

- **(infinite-dimensional) Lie theory** (Neeb, 2005) – lecture notes; (Neeb, 2006) – extensive survey; and, once it becomes available (Glöckner and Neeb, forthcoming).
- **Manifolds of mappings** (Wockel, 2014); on non-compact domains (Michor, 1980).
- **(weak) Riemannian geometry** (Bruveris, 2018, 2019).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Furthermore, there is a host of topics that could not be covered in this book. Albeit there is again no hope that we could do justice to the vast body of research on these topics, we would like to mention a few works that are either introductory or exhibit new phenomena which are genuinely infinitely dimensional in nature:

- **Symplectic geometry** (Abraham et al., 1988 (Banach setting); Kriegl and Michor (1997, Section 48)).
- **Sub-Riemannian geometry** (Grong et al., 2015; Agrachev and Caponigro, 2009).
- **Kähler geometry** (Sergeev, 2020).
- **Poisson geometry** (Beltiţă et al., 2018).
- **Finsler geometry** (Larotonda (2019) on diffeomorphism groups; Eftekharinasab and Petrusenko (2020) in the context of bounded Fréchet geometry).

Acknowledgements

We acknowledge support by Nord University and the Trond Mohn foundation in the 'Pure Mathematics in Norway' program, see www.puremath.no/ for more information. In particular, I am thankful for Nord University's commitment to making the book available in open access format.

I have profited from an early draft of the forthcoming book by H. Glöckner and K.-H. Neeb on infinite-dimensional Lie groups (Glöckner and Neeb, forthcoming). Their lucid presentation of calculus directly influenced §1. Material from the thesis by A. Mykleblust ('Limits of Banach spaces', UiB 2020, advised by A. Schmeding) is included in Appendix A.

I am indebted to Martin Bauer, Rafael Dahmen, Gard O. Helle, Cy Maor, Klas Modin, Torstein Nilssen and David M. Roberts for many useful suggestions and discussions while I was writing this book. Further thanks go to Rafael Dahmen, Klas Modin, Torstein Nilssen, David Prinz and Nikolas Tapia for reading the manuscript and suggesting many improvements. Finally, I thank the participants of the course MAT331 who brought many errors in previous versions to my attention.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

1

Calculus in Locally Convex Spaces

1.1 Introduction

It is well known that 'multidimensional calculus', aka 'Fréchet calculus', carries over to the realm of Banach spaces and Banach manifolds (see e.g. Lang, 1999). As we have seen in the Preface, Banach spaces are often not sufficient for our purposes. To generalise derivatives we will, as a minimum, need vector spaces with an amenable topology (which need not be induced by a norm).

1.1 Definition Consider a vector space E. A topology $\mathcal{T}$ on E making addition $+\colon E \times E \to E$ and scalar multiplication $\cdot\colon \mathbb{R} \times E \to E$ continuous is called a *vector topology* (where $\mathbb{R}$ carries the usual norm topology). We then say that $(E, \mathcal{T})$ (or E for short) is a *topological vector space* (or *TVS* for short).

1.2 Example (a) Every normed space and, in particular, every finite-dimensional vector space is a topological vector space.

(b) For a more interesting example, fix $0 < p < 1$. Two measurable functions $\gamma, \eta\colon [0,1] \to \mathbb{R}$ are equivalent $\gamma \sim \eta$ if and only if $\int_0^1 |\gamma(s) - \eta(s)| \mathrm{d}s = 0$. Denote by $L^p[0,1]$ the vector space of all equivalence classes $[\gamma]$ of functions such that $\int_0^1 |\gamma(s)|^p \mathrm{d}s < \infty$. Topologise $L^p[0,1]$ via the metric topology induced by

$$d([\gamma],[\eta]) \coloneqq \int_0^1 |\gamma(s) - \eta(s)|^p \mathrm{d}s.$$

In a metric space, we can test continuity of the vector space operations using sequences. For this, pick $\lambda_n \to \lambda \in \mathbb{R}$ and $[\gamma_n] \to [\gamma], [\eta_n] \to [\eta]$ (with respect to d) and use the triangle inequality to obtain:

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$d(\lambda_n[\gamma_n] + [\eta_n], \lambda[\gamma] + [\eta]) \\ \leq |\lambda_n - \lambda|^p d([\gamma_n],[0]) + |\lambda|^p d([\gamma_n],[\gamma]) + d([\eta_n],[\eta]).$$

This shows that the vector space operations are continuous, that is, $L^p[0,1]$ is a TVS.

In topological vector spaces, differentiable curves can be defined as follows:

1.3 Definition Let E be a topological vector space. A continuous mapping $\gamma\colon I \to E$ from a non-degenerate interval[1] $I \subseteq \mathbb{R}$ is called a C^0-curve. A C^0-curve is called a C^1-*curve* if the limit

$$\gamma'(s) := \lim_{t\to 0} \frac{1}{t}(\gamma(s+t) - \gamma(s))$$

exists for all $s \in I^\circ$ (interior of I) and extends to a continuous map $\frac{\mathrm{d}}{\mathrm{d}t}\gamma := \gamma'\colon I \to E$, $s \mapsto \gamma'(s)$. Recursively for $k \in \mathbb{N}$, we call γ a C^k-*curve* if γ is a C^{k-1}-curve and $\frac{\mathrm{d}^{k-1}}{\mathrm{d}t^{k-1}}\gamma$ is a C^1-curve. Then $\frac{\mathrm{d}^k}{\mathrm{d}t^k}\gamma := \left(\frac{\mathrm{d}^{k-1}}{\mathrm{d}t^{k-1}}\gamma\right)'$. If γ is a C^k-curve for every $k \in \mathbb{N}_0$, we also say that γ is *smooth* or of class C^∞.

Unfortunately, calculus on topological vector spaces is, in general, ill behaved. The next exercise shows that derivatives may fail to give us meaningful information.

Exercises

1.1.1 Given $0 < p < 1$ we let $L^p[0,1]$ be the topological vector space from Example 1.2(b). Recall that the topology on $L^p[0,1]$ is induced by the metric $d([\gamma],[\eta]) := \int_0^1 |\gamma(s) - \eta(s)|^p \mathrm{d}s$. For a set $A \subseteq [0,1]$ write $\mathbf{1}_A$ for the characteristic function and define

$$\beta\colon [0,1] \to L^p[0,1], \quad \beta(t) := [\mathbf{1}_{[0,t[}].$$

Show that β is an injective C^1-curve with $\beta'(t) = 0$, for all $t \in [0,1]$.

Obviously we would like to avoid this defect, and so we have to strengthen the assumptions on our vector spaces.

[1] That is, I has more than one point. In the following, we will always assume this when talking about intervals.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

1.2 Curves in Locally Convex Spaces

Calculus in topological vector spaces exhibits pathologies that can be avoided by strengthening the requirements on the underlying space. This leads to locally convex spaces, whose topology is induced by so-called seminorms. See also Appendix A for more information on locally convex spaces.

1.4 Definition Let E be a vector space. A map $p\colon E \to [0,\infty[$ is called a *seminorm* if it satisfies the following:

(a) $p(\lambda x) = |\lambda| p(x), \forall \lambda \in \mathbb{R}, x \in E$,
(b) $p(x+y) \leq p(x) + p(y)$.

Note that, in contrast with the definition of a norm, we did not require that $p(x) = 0$ if and only if $x = 0$. The next definition uses the notion of an initial topology, which we recall for the reader's convenience in Appendix B.

1.5 Definition A topological vector space $(E,\mathcal{T})$ is called a *locally convex space* if there is a family $\{p_i : E \to [0,\infty[\mid i \in I\}$ of continuous seminorms for some index set I such that

(a) $\mathcal{T}$ is the initial topology with respect to the canonical projections $\{q_i : E \to E/p_i^{-1}(0)\}_{i\in I}$ onto the normed spaces $E/p_i^{-1}(0)$.
(b) If $x \in E$ with $p_i(x) = 0$ for all $i \in I$, then $x = 0$. Thus the seminorms separate the points, that is, $\mathcal{T}$ has the Hausdorff property.[2]

We then say that the topology $\mathcal{T}$ is *generated by the family of seminorms* $\{p_i\}_{i\in I}$ and call this family a *generating family of seminorms*. Usually we suppress $\mathcal{T}$ and write $(E,\{p_i\}_{i\in I})$ or simply E instead of $(E,\mathcal{T})$.

Alternative to (a) We will see in Appendix A that equivalent to (a), we can define $\mathcal{T}$ to be the unique vector topology determined by the basis of 0-neighbourhoods given by (finite) intersections of the balls $B_{i,\varepsilon}(0) = \{x \in E \mid p_i(x) < \varepsilon\}$, where p_i runs through a generating family of seminorms. These balls are all convex, thus justifying the name locally convex space.

A locally convex space $(E,\{p_i\}_{i\in\mathbb{N}})$ with a countable system of seminorms is *metrisable* (i.e. its topology is induced by a metric; see Exercise 1.2.1) and if E is complete, it is called *Fréchet space*.

1.6 Example (a) Every normed space $(E,\|\cdot\|)$ is a locally convex space, where the family of seminorms consists only of the norm $\|\cdot\|$.

[2] Some authors do not require separation of points, whence our locally convex spaces are Hausdorff locally convex spaces in their terminology.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(b) Consider the space $C^\infty([0,1],\mathbb{R})$ of all smooth functions from the interval $[0,1]$ to $\mathbb{R}$ (with pointwise addition and scalar multiplication). This space is not naturally a normed space.[3] We define a family of seminorms on it via

$$\|f\|_n := \sup_{0\le k\le n} \left\| \frac{\mathrm{d}^k}{\mathrm{d}t^k} f \right\|_\infty = \sup_{0\le k\le n} \sup_{t\in[0,1]} \left| \frac{\mathrm{d}^k}{\mathrm{d}t^k} f(t) \right|, n \in \mathbb{N}_0.$$

The topology generated by the seminorms is called the *compact-open C^∞-topology* and turns $C^\infty([0,1],\mathbb{R})$ into a locally convex space, which is even a Fréchet space (Exercise 1.2.2).

Locally convex spaces have many good properties, for example, they admit enough continuous linear functions to separate the points, that is, the following holds.

1.7 Theorem (Hahn–Banach (Meise and Vogt, 1997, Proposition 22.12)) *For a locally convex space E the continuous linear functionals separate the points, that is, for each pair $x, y \in E$ there exists a continuous linear $\lambda\colon E \to \mathbb{R}$ such that $\lambda(x) \neq \lambda(y)$.*

1.8 Definition Let E be a locally convex space, then we denote by $E' = L(E,\mathbb{R})$ the continuous linear maps from E to $\mathbb{R}$. The space E' is the so-called *dual space* of E. There are several ways to turn E' into a locally convex space (Rudin, 1991, p. 63f) but, in general, we will not need a topology beyond the special case if E is a Banach space and E' carries the operator norm topology.

With the help of the Hahn–Banach theorem, we can avoid the pathologies observed for topological vector spaces. To this end, we need the notion of a weak integral.

1.9 Definition Let $\gamma\colon I \to E$ be a C^0-curve in a locally convex space E and $a, b \in I$. If there exists $z \in E$ such that

$$\lambda(z) = \int_a^b \lambda(\gamma(t))\mathrm{d}t, \quad \forall \lambda \in E',$$

then $z \in E$ is called the *weak integral of* γ from a to b and denoted $\int_a^b \gamma(t)\mathrm{d}t := z$.

Note that weak integrals (if they exist) are uniquely determined due to the Hahn–Banach theorem.

[3] For any normed topology, the differential operator $D\colon C^\infty([0,1],\mathbb{R}) \to C^\infty([0,1],\mathbb{R})$, $D(f) = f'$ must be discontinuous (which is certainly undesirable). To see this, recall that a continuous linear map on a normed space has bounded spectrum, but D has arbitrarily large eigenvalues (consider $f_n(t) := \exp(nt)$, $n \in \mathbb{N}$).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

1.10 Proposition (First part of the fundamental theorem of calculus) *Let $\gamma\colon I \to E$ be a C^1-curve in a locally convex space E and $a, b \in I$, then*

$$\gamma(b) - \gamma(a) = \int_a^b \gamma'(t)\mathrm{d}t.$$

Proof Let $\lambda \in E'$. It is easy to see that $\lambda \circ \gamma\colon I \to \mathbb{R}$ is a C^1-curve with $(\lambda \circ \gamma)' = \lambda \circ (\gamma')$. The standard fundamental theorem of calculus yields

$$\lambda(\gamma(b) - \gamma(a)) = \lambda(\gamma(b)) - \lambda(\gamma(a)) = \int_a^b (\lambda \circ \gamma)'(s)\mathrm{d}s = \int_a^b \lambda(\gamma'(s))\mathrm{d}s.$$

Hence $z = \gamma(b) - \gamma(a)$ satisfies the defining property of the weak integral. □

Note that Proposition 1.10 implies that $L^p[0,1]$ cannot be a locally convex space for $0 < p < 1$; see Rudin (1991, 1.47) for an elementary proof of this fact.

1.11 Remark Also the second part of the fundamental theorem of calculus is true in our setting. Thus if $\gamma\colon I \to E$ is a C^0-curve, $a \in I$ and the weak integral

$$\eta(t) := \int_a^t \gamma(s)\mathrm{d}s$$

exists for all $t \in I$. Then $\eta\colon I \to E$ is a C^1-curve in E, and $\eta' = \gamma$.

The proof, however, needs more techniques based on convex sets which we do not wish to go into (see Glöckner and Neeb, forthcoming).

The reader may wonder now, when do weak integrals of curves exist? One can prove that weak integrals of continuous curves always exist *in the completion of a locally convex space*. The key point is that the integrals can be defined using Riemann sums, but these do not necessarily converge in the space itself (Kriegl and Michor, 1997, Lemma 2.5). Thus weak integrals exist for suitably complete spaces. To avoid getting bogged down with the discussion of completeness properties, we define the following:

1.12 Definition A locally convex space E is *Mackey complete* if for each smooth curve $\gamma\colon [0,1] \to E$ there exists a smooth curve $\eta\colon [0,1] \to E$ with $\eta' = \gamma$.

Due to the fundamental theorem of calculus this implies that $\eta(s) - \eta(0) = \int_0^s \gamma(t)\mathrm{d}t$. Thus the weak integral of smooth curves exists in Mackey complete spaces.

1.13 Remark Mackey completeness is a very weak completeness condition, in particular, sequential completeness (i.e. Cauchy sequences converge in the space) implies Mackey completeness. This is evident from the

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

alternative characterisation of Mackey completeness using sequences; see Definition A.1. Note, however, that it is not entirely trivial to find examples of Mackey complete but not sequentially complete spaces. We mention here that the space $\mathcal{K}(E,F)$ of compact operators between two (infinite-dimensional) Banach spaces E,F with the strong operator topology is not sequentially complete but Mackey complete (see Voigt, 1992).

However, in metrisable locally convex spaces (e.g. in normed spaces) Mackey completeness is equivalent to completeness; see Jarchow (1981, 10.1.4). We refer to Kriegl and Michor (1997, I.2) for more information on Mackey completeness. In particular, Kriegl and Michor (1997, Theorem 2.14) show that integrals exist for C^1-curves in Mackey complete spaces.

So far we have defined differentiable curves with values in locally convex spaces. The next step is to consider differentiable mappings between locally convex spaces. Here a different notion of calculus is needed. It turns out that (even on Fréchet spaces) there are many generalisations of Fréchet calculus (see Keller, 1974) without a uniquely preferable choice. In the next section, we present a simple and versatile notion called Bastiani calculus. Another popular approach to calculus in locally convex spaces, the so-called convenient calculus, is discussed in Appendix A.7.

Exercises

1.2.1 Let $(E,\{p_n\}_{n\in\mathbb{N}})$ be a locally convex space whose topology is generated by a countable set of seminorms. Prove that

$$d(x,y) := \sum_{n\in\mathbb{N}} 2^{-n}\,\frac{p_n(x-y)}{p_n(x-y)+1}$$

is a metric on E and the metric topology coincides with the locally convex topology.

1.2.2 Consider $C^\infty([0,1],\mathbb{R})$ with the compact open C^∞-topology (see Example 1.6).

(a) Show that a sequence $(f_k)_{k\in\mathbb{N}}$ converges to f in this topology if and only if for all $\ell\in\mathbb{N}_0$ $\left(\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell}f_k\right)_k$ converges uniformly to $\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell}f$.
Hint: The uniform limit of a sequence of continuous functions is continuous. If a function sequence and the sequence of (first) derivatives converges, the limit of the sequence is differentiable.

(b) Deduce that every Cauchy sequence in the compact open C^∞-topology converges to a smooth function. As $C^\infty([0,1],\mathbb{R})$ is

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

a metric space by Exercise 1.2.1, this implies that the space is complete, that is, a Fréchet space.

(c) Show that the differential operator

$$D\colon C^\infty([0,1],\mathbb{R}) \to C^\infty([0,1],\mathbb{R}), \qquad f \mapsto f'$$

is continuous linear. *Hint:* Lemma A.5.

1.2.3 Let $(E,\{p_i\}_I)$ be a locally convex space whose topology is generated by a *finite* set of seminorms. Show that $p(x) = \max_{i\in I} p_i(x)$ defines a norm on E, which induces the same topology as the family $\{p_i\}$. In this case we call E normable.

1.2.4 Establish the following properties of weak integrals:

(a) If the weak integrals of $\gamma,\eta\colon [a,b] \to E$ from a to b exist and $s \in \mathbb{R}$, then also the weak integral of $\gamma + s\eta$ exists and $\int_a^b (\gamma(t) + s\eta(t))\mathrm{d}t = \int_a^b \gamma(t)\mathrm{d}t + s\int_a^b \eta(t)\mathrm{d}t$.

(b) If $\gamma\colon [a,b] \to E$ is constant, $\gamma(t) \equiv K$, then $\int_a^b \gamma(t)\mathrm{d}t$ exists and equals $(b-a)K$.

(c) $\int_a^c \gamma(t)\mathrm{d}t = \int_a^b \gamma(t)\mathrm{d}t + \int_b^c \gamma(t)\mathrm{d}t$ (if the integrals exist).

1.2.5 Let $\gamma\colon I \to E$ be a C^k-curve ($k \in \mathbb{N}$) and $\lambda\colon E \to F$ be continuous linear for E,F locally convex. Show that $\lambda\circ\gamma$ is C^k such that $\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell}(\lambda\circ\gamma) = \lambda\circ\left(\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell}\gamma\right)$, $1 \le \ell \le k$.

1.2.6 Endow a vector space E with a topology $\mathcal{T}$ generated by seminorms as in Definition 1.5. Show that $(E,\mathcal{T})$ is a topological vector space (and so requiring that locally convex spaces are topological vector spaces was superfluous).

1.3 Bastiani Calculus

Bastiani calculus (also called Keller's C_c^k-theory; Keller, 1974), introduced in Bastiani (1964), builds a calculus around directional derivatives and their continuity. It is the basis of our investigation as this calculus works in locally convex spaces beyond the Banach setting.

1.14 Definition Let E,F be locally convex spaces, $U \subseteq E$, $f\colon U \to F$ a map and $r \in \mathbb{N}_0 \cup \{\infty\}$. If it exists, we define for $(x,h) \in U \times E$ the *directional derivative*

$$df(x;h) := D_h f(x) := \lim_{\mathbb{R}\setminus\{0\}\ni t\to 0} t^{-1}\big(f(x+th) - f(x)\big).$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

We say that f is C^r if the *iterated directional derivatives*

$$d^k f(x; y_1, \dots, y_k) := (D_{y_k} D_{y_{k-1}} \cdots D_{y_1} f)(x)$$

exist for all $k \in \mathbb{N}_0$ such that $k \leq r$, $x \in U$ and $y_1, \dots, y_k \in E$ and define continuous maps $d^k f : U \times E^k \to F$ (where $d^0 f := f$). If f is C^k for all $k \in \mathbb{N}_0$ we say that f is *smooth* or C^∞. Note that $df = d^1 f$ and for curves $c : I \to E$ we have $c'(t) = dc(t; 1)$.

1.15 Remark Note that the iterated directional derivatives are only taken with respect to the first variable (i.e. of the map $x \mapsto df(x; v)$, where v is supposed to be fixed). One can alternatively define iterated differentials to derivate with respect to all variables, but this leads to the same differentiability concept (see Glöckner, 2002 for a detailed explanation). The following observations are easily proved from the definitions:

(a) $d^2 f(x; v, w) = \lim_{t \to 0} t^{-1}(df(x + tw; v) - df(x; v))$.
(b) $d^k f(x; v_1, \dots, v_k) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} d^{k-1} f(x + tv_k; v_1, \dots, v_{k-1})$.
(c) f is C^k if and only if f is C^{k-1} and $d^{k-1} f$ is C^1. Then $d^k f = d(d^{k-1} f)$.

Finally, there is a version of the *Schwarz theorem* which states that the order of directions $v_1, \dots, v_k$ in $d^k f(x; v_1, \dots, v_k)$ is irrelevant (see Exercise 1.3.3).

1.16 Example Let $A : E \to F$ be a continuous linear map between locally convex spaces. Then A is C^1, as we can exploit

$$dA(x; v) = \lim_{t \to 0} t^{-1}(A(x + tv) - A(x)) = \lim_{t \to 0} A(v) = A(v).$$

In particular, since A is continuous, so is the first derivative and we see that A is a C^1-map. Computing the second derivative, we use that the first derivative is constant in x (but not in v!) to obtain

$$\begin{aligned} d^2 A(x; v, w) =& D_w(dA(x; v)) = \lim_{t \to 0} t^{-1}(dA(x + tw; v) - dA(x; v)) \\ =& \lim_{t \to 0} t^{-1}(A(v) - A(v)) = 0. \end{aligned}$$

In conclusion A is a C^2-map (obviously even a C^∞-map) whose higher derivatives vanish.

1.17 Lemma *Let $f : E \supseteq U \to F$ be a C^1-map. Then $df(x; \cdot)$ is homogeneous, that is, $df(x; sv) = s\,df(x; v)$ for all $x \in U, v \in E$ and $s \in \mathbb{R}$.*

Proof As $df(x; 0v) = df(x; 0) = 0 = 0\,df(x; v)$, we may assume that $s \neq 0$ and thus $df(x; sv) = \lim_{t \to 0} t^{-1}(f(x + tsv) - f(x)) = s \lim_{t \to 0} (st)^{-1}(f(x + tsv) - f(x)) = s\,df(x; v)$. □

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

1.18 Proposition (Mean value theorem on locally convex spaces) *Let E, F be locally convex spaces and $f: U \to F$ a C^1-map on $U \subseteq E$. Then*

$$f(y) - f(x) = \int_0^1 df(x + t(y-x); y-x)\mathrm{d}t \tag{1.1}$$

for all $x, y \in U$ such that U contains the line segment $\overline{xy} := \{tx + (1-t)y \mid t \in [0,1]\}$.

Proof Note that the curve $\gamma: [0,1] \to F, \gamma(t) := f(x + t(y-x))$ is differentiable at each $t \in [0,1]$. Its derivative is

$$\gamma'(t) = \lim_{s \to 0} s^{-1}(\gamma(t+s) - \gamma(t)) = df(x + t(y-x), y-x),$$

whence γ' is continuous (as df is) and thus a C^1-curve. Apply now the Fundamental theorem 1.10 to γ' to obtain (1.1). □

On a locally convex space, every point has arbitrarily small convex neighbourhoods. Convex neighbourhoods contain all line segments between points in the neighbourhood, whence Proposition 1.18 is available on these neighbourhoods. As a consequence we obtain the following.

1.19 Corollary *If $f: U \to F$ is a C^1-map with $df \equiv 0$, then f is locally constant.*

Proof For $x \in U$ choose a convex neighbourhood $x \in V \subseteq U$ (see Appendix A). For each $y \in V$ the line segment connecting x and y is contained in V, and so the vanishing of the derivative with (1.1) implies $f(x) = f(y)$ and f is constant on V. □

1.20 Proposition (Rule on partial differentials) *Let E_1, E_2, F be locally convex spaces, $U \subseteq E_1 \times E_2$ and let $f: U \to F$ be continuous. Then f is C^1 if and only if the limits*

$$d_1 f(x,y;v_1) := \lim_{t \to 0} t^{-1}(f(x + tv_1, y) - f(x,y)),$$
$$d_2 f(x,y;v_2) := \lim_{t \to 0} t^{-1}(f(x, y + tv_2) - f(x,y))$$

exist for all $(x,y) \in U$ and $(v_1, v_2) \in E_1 \times E_2$ and extend to continuous mappings $d_i f: U \times E_i \to F, i = 1,2$. In this case,

$$df(x,y;v_1,v_2) = d_1 f(x,y;v_1) + d_2 f(x,y;v_2), \quad \forall (x,y) \in U, (v_1,v_2) \in E_1 \times E_2. \tag{1.2}$$

Proof If f is C^1 the mappings $d_i f$ clearly exist and are continuous. Conversely, let us assume that the mappings $d_i f$ exist and are continuous. For

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$(x,y) \in U$, $(v_1,v_2) \in E \times F$, we fix $\varepsilon > 0$ such that $(x,y) + t(v_1,v_2) \in U$ whenever $|t| < \varepsilon$. Now if we fix the ith component of f we obtain a C^1-mapping (by hypothesis, the derivative is $d_i f$). Therefore Proposition 1.18 with Lemma 1.17 yields

$$\begin{aligned}
&\frac{f((x,y)+t(v_1,v_2)) - f(x,y)}{t} \\
&= \frac{f(x+tv_1, y+tv_2) - f(x+tv_1, y)}{t} + \frac{f(x+tv_1,y) - f(x,y)}{t} \\
&= \int_0^1 d_2 f(x+tv_1, y+stv_2; v_2)\mathrm{d}s + \int_0^1 d_1 f(x+stv_1, y; v_1)\mathrm{d}s. \qquad (1.3)
\end{aligned}$$

The integrals (1.3) make sense also for $t = 0$, whence they define maps $I_i :]-\varepsilon,\varepsilon[\to H$. Due to continuous dependence on the parameter t,[4] the right-hand side of (1.3) converges for $t \to 0$. We deduce that the limit df exists and satisfies (1.2) which is continuous, whence f is C^1. □

The following alternative characterisation of C^1-maps will turn the proof of the chain rule into a triviality. However, we shall only sketch the proof to avoid discussing convergence issues of the weak integral involved.

1.21 Lemma *A map $f : E \supseteq U \to F$ is of class C^1 if and only if there exists a continuous mapping, the difference quotient map,*

$$f^{[1]} : U^{[1]} := \{(x,v,s) \in U \times E \times \mathbb{R} \mid x + sv \in U\} \to F$$

such that $f(x+sv) - f(x) = s f^{[1]}(x,v,s)$ for all $(x,v,s) \in U^{[1]}$.

Proof Let us assume first that $f^{[1]}$ exists and is continuous. Note that $U^{[1]} \subseteq U \times E \times \mathbb{R}$. Then $df(x;v) = f^{[1]}(x,v,0)$ exists and is continuous as a partial map of $f^{[1]}$. So f is C^1. Conversely, if f is C^1, the map

$$f^{[1]}(x,v,s) := \begin{cases} s^{-1}(f(x+sv) - f(x)), & (x,v,s) \in U^{[1]}, s \neq 0, \\ df(x;v), & (x,v,s) \in U^{[1]}, s = 0 \end{cases}$$

is continuous on the open set $U^{[1]} \setminus \{(x,v,s) \in U^{[1]} \mid s = 0\}$. That $f^{[1]}$ extends to a continuous map on all of $U^{[1]}$ follows from continuity of parameter-dependent weak integrals; see Bertram et al. (2004, Proposition 7.4) for details. □

1.22 Lemma *If $f : E \supseteq U \to F$ is C^1, then $df(x;\cdot) : E \to F$ is a continuous linear map for each $x \in U$.*

[4] We are cheating here; the continuous dependence of weak integrals on parameters has not been established in this book. See Hamilton (1982, I Theorem 2.1.5) for a proof that carries over to our setting.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Proof Fix $x \in U$ and note that $df(x;\cdot)$ is continuous as a partial map of the continuous df. We have already seen in Lemma 1.17 that $df(x;sv) = sdf(x;v)$ for all $s \in \mathbb{R}$, and so we only have to prove additivity. Choosing $v, w \in E$ we compute

$$\begin{aligned} t^{-1}&(f(x+t(v+w)) - f(x)) \\ &= t^{-1}(f(x+t(v+w)) - f(x+tv)) + t^{-1}(f(x+tv) - f(x)) \\ &= f^{[1]}(x+tv, w, t) + f^{[1]}(x, v, t). \end{aligned}$$

The right-hand side also makes sense for $t = 0$ and is continuous, whence passing to the limit we get $df(x; v+w) = df(x;v) + df(x;w)$. □

1.23 Proposition (Chain rule) *Let $f\colon E \supseteq U \to F$ and $g\colon F \supseteq W \to K$ be C^1-maps with $f(U) \subseteq W$. Then $g \circ f$ is a C^1-map with derivative given by*

$$d(g\circ f)(x;v) = dg(f(x); df(x,v)) \quad (\text{i.e.} (g\circ f)'(x) = g'(f(x)) \circ f'(x)).$$

Proof We use the notation from Lemma 1.21 and write for $(x,y,t) \in U^{[1]}$ with $t \neq 0$,

$$\begin{aligned} t^{-1}(g(f(x+ty)) - g(f(x))) &= t^{-1}\left(g\left(f(x) + t\frac{f(x+ty)-f(x)}{t}\right) - g(f(x))\right) \\ &= g^{[1]}(f(x), f^{[1]}(x,y,t), t). \end{aligned} \tag{1.4}$$

The function $h\colon U^{[1]} \to K, h(x,y,t) := g^{[1]}(f(x), f^{[1]}(x,y,t), t)$ is continuous and extends the right-hand side of (1.4). Hence Lemma 1.21 shows that $g \circ f$ is C^1, with

$$(g\circ f)^{[1]}(x,y,t) = g^{[1]}(f(x), f^{[1]}(x,y,t), t) \text{ for all } (x,y,t) \in U^{[1]}.$$

Thus

$$\begin{aligned} d(g\circ f)(x;y) &= (g\circ f)^{[1]}(x,y,0) = g^{[1]}(f(x), f^{[1]}(x,y,0), 0) \\ &= dg(f(x); df(x;y)). \end{aligned}$$ □

The chain rule is the basis to transport concepts from differential geometry such as manifolds, tangent spaces and so on to our setting. Later chapters will define these objects.

1.24 Lemma *Let $E, (F_i)_{i\in I}$ be locally convex spaces and $f\colon E \supseteq U \to \prod_{i\in I} F_i$ a map on an open subset. Let $k \in \mathbb{N}_0 \cup \{\infty\}$ and set $f_i\colon \mathrm{pr}_i \circ f\colon U \to F_i$ (where pr_i is the ith* coordinate projection*). Then f is C^k if and only if every $f_i, i \in I$ is C^k and*

$$df(x;v) = (df_i(x;v))_{i\in I}, \quad x \in U, v \in E. \tag{1.5}$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Proof If f is C^k, we note that every $f_i = \mathrm{pr}_i \circ f$ is C^k by the chain rule as the projections are continuous linear, hence smooth. Further, $df_i = \mathrm{pr}_i \circ df$ which establishes (1.5).

For the converse, note that it suffices to assume that $k \in \mathbb{N}$. We argue by induction starting with $k = 1$. Then $t^{-1}f(x+tv) - f(x)) = (t^{-1}(f_i(x+tv) - f_i(x))_{i\in I}$. Now the components converge to $df_i(x;v)$ for $t \to 0$, whence the difference quotient converges to the limit in (1.5). Since every df_i is continuous, $df = (df_i)_{i\in I} : U \times E \to \prod_{i\in I} F_i$ is also continuous and f is C^1. For the induction step we notice that by induction, $\mathrm{pr}_i \circ df = df_i$ is C^{k-1}. By induction, df is C^{k-1} and thus f is C^k. □

A subset $Y \subseteq X$ of a topological space X is called *sequentially closed* if $\lim_{n\to\infty} x_n \in A$ for each sequence $(x_n)_{\mathbb{N}} \subseteq A$ which converges in X. The following lemma will be useful in the discussion of submanifolds in Section 1.5.

1.25 Lemma *Let $f : E \supseteq U \to F$ be a continuous map from an open subset of a locally convex space to a locally convex space, and $F_0 \subseteq F$ a sequentially closed vector subspace such that $f(U) \subseteq F_0$. Let $k \in \mathbb{N} \cup \{\infty\}$, then f is C^k if and only if the corestriction $f|^{F_0} : U \to F_0$ is C^k.*

Proof If $f|^{F_0}$ is C^k, so is $f = \iota \circ f|^{F_0}$, where $\iota : F_0 \to F$ is the continuous linear (hence smooth) inclusion.

Conversely, we argue by induction and assume first that f is C^1. For $x \in U$ and $v \in E$, pick a sequence $t_n \to 0$ such that $x + t_n v \in U$ for each $n \in \mathbb{N}$. Then $df(x;y) = \lim_{n\to\infty}(f(x+t_n v) - f(x)) \in F_0$ by sequential closedness. Hence the limit exists in F_0. Further, as a map $U \times E \to F_0$, $(x,v) \mapsto df(x;v)$ is continuous. We conclude that $f|^{F_0}$ is C^1. If f is C^k, $d(f|^{F_0}) = (df)|^{f_0}$ is C^{k-1} by induction and hence f is C^k. □

In Example A.32 we will see that in Lemma 1.25 sequential closedness is a necessary assumption for the validity of the statement.

Exercises

1.3.1 Check the details for Remark 1.15 (for Schwarz' theorem see below).

1.3.2 Let E_1, E_2, F be locally convex spaces and $\beta : E_1 \times E_2 \to F$ be a continuous bilinear map. Show that β is C^1 with first derivative $d\beta(x_1,x_2;y_1,y_2) = \beta(x_1,y_2) + \beta(y_1,x_2)$. Compute all higher derivatives of β and show that β is smooth.

1.3.3 *Schwarz Theorem*: If $f : E \supseteq U \to F$ is a C^k-map, and $x \in U$, prove that $d^r f(x;\cdot) : E^r \to F$ is symmetric for all $2 \le r \le k$ (i.e. the order of arguments is irrelevant to the function value).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Hint: By the Hahn–Banach theorem, it suffices to consider $\lambda \circ d^r f(x;\cdot) = d^r \lambda \circ f(x;\cdot)$, whence without loss of generality $F = \mathbb{R}$. Now for fixed $v_1,\dots,v_r$, the property can be checked in the finite-dimensional subspace generated by $v_1,\dots,v_r$.

1.3.4 The continuous functions $C([0,1],\mathbb{R})$ form a Banach space with respect to the supremum norm $\|\cdot\|_\infty$ (the resulting topology is the compact open topology).

(a) Let $f\colon \mathbb{R} \to \mathbb{R}$ be continuous and $(\gamma_n)_{n\in\mathbb{N}} \subseteq C([0,1],\mathbb{R})$ be a uniformly convergent sequence of functions with limit γ. Exploit that f is uniformly continuous on each ball and show that $f\circ\gamma_n \to f\circ\gamma$ uniformly. Deduce that the pushforward $f_*\colon C([0,1],\mathbb{R}) \to C([0,1],\mathbb{R}), \eta \mapsto f\circ\eta$ is continuous

Assume that $f \in C^1(\mathbb{R},\mathbb{R})$. Our aim will be to see that f_* is then C^1.

(b) Assume that the limit

$$df_*(\gamma;\eta) := \lim_{t\to 0} t^{-1}(f_*(\gamma+t\eta) - f_*(\gamma)) \tag{1.6}$$

exists. The point evaluation $\mathrm{ev}_x\colon C([0,1],\mathbb{R}) \to \mathbb{R}, \eta \mapsto \eta(x)$ is continuous linear for each $x \in [0,1]$. Apply ev_x to both sides of (1.6) and find the only possible candidate $\psi(\gamma,\eta)$ for $df_*(\gamma;\eta)$.

(c) Use point evaluations to verify that

$$t^{-1}(f_*(\gamma+t\eta) - f_*(\gamma)) = \int_0^1 \psi(\gamma+st\eta,\eta)\mathrm{d}s.$$

(d) Show that $\psi(\gamma,\eta)$ from (b) is indeed the directional derivative $df_*(\gamma;\eta)$.

(e) Verify that df_* is continuous, hence f_* is C^1.

1.3.5 Let E,F,H be locally convex spaces, $U \subseteq E, V \subseteq H$, $f\colon U \to F$ a C^2-map and let $g\colon V \to U$ and $h\colon V \to E$ be C^1. Prove that the differential of the C^1-map $\phi := df\circ(g,h)\colon V \to F, \phi(x) = df(g(x);h(x))$ is given by

$$d\phi(x;y) = d^2 f(g(x);h(x),dg(x;y)) + df(g(x);dh(x;y)), \tag{1.7}$$

for all $x \in V$, $y \in H$.

1.4 Bastiani versus Fréchet Calculus on Banach Spaces

On Banach spaces one usually defines differentiability in terms of the so-called Fréchet derivative. We briefly recall the definitions that should be familiar from

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

basic courses on calculus. While Bastiani calculus is somewhat weaker than Fréchet calculus, the gap between those two can be quantified. We collect these results in the present section.

1.26 Definition A map $f\colon E \supseteq U \to F$ from an open subset of a normed space $(E, \|\cdot\|_E)$ to a normed space $(F, \|\cdot\|_F)$ is *continuous Fréchet differentiable* (or FC^1) if, for each $x \in U$, there exists a continuous linear map $A_x \in L(E,F)$ such that

$$f(x+h) - f(x) = A_x \cdot h + R_x(h) \text{ with } \lim_{h\to 0} \|R_x(h)\|_F / \|h\|_E = 0$$

and the mapping $Df\colon U \to L(E,F)$, $x \mapsto A_x$ is continuous (where the right-hand side carries the operator norm). Inductively we define for $k \in \mathbb{N}$ that f is a *k-times continuous Fréchet differentiable* map (or FC^k-map) if it is FC^1 and Df is FC^{k-1}. Moreover, f is (Fréchet-)smooth or FC^∞ if f is an FC^k map for every $k \in \mathbb{N}$.

The reader may wonder why the notion of Fréchet differentiability cannot be generalised beyond the setting of normed spaces. The reason for this is that the continuity of the derivative cannot be formulated as there is no suitable topology on the spaces $L(E,F)$. Indeed, there is no locally convex topology making evaluation and composition on the spaces $L(E,F)$ continuous. We refer to Proposition A.19 for an example of this pathology in the context of dual spaces.

From the definition, it is apparent (Exercise 1.4.2) that if f is FC^1, then $Df(x)(h) = df(x;h)$ and thus every FC^1-map is automatically C^1 in the Bastiani sense. However, one learns in basic calculus courses that existence of directional derivatives is weaker than the existence of derivatives in the Fréchet sense. The next example exhibits this.

1.27 Example (Bastiani C^1 is weaker than Fréchet C^1 (Milnor 1982)) Consider the Banach space

$$\ell^1 = \left\{ (x_n)_{n\in\mathbb{N}} \mid x_n \in \mathbb{R}, \text{ for all } n \in \mathbb{N}, \|(x_n)\|_{\ell^1} = \sum_{n\in\mathbb{N}} |x_n| < \infty \right\}.$$

Let $\varphi(u) := \log(1+u^2)$ and $\psi(u) := \frac{\mathrm{d}}{\mathrm{d}u}\varphi(u) = \frac{2u}{1+u^2}$. We observe that $|\psi(u)| \leq 1$, whence $|\varphi(u)| \leq |u|$ and we obtain a well-defined map

$$f\colon \ell^1 \to \mathbb{R}, f((x_n)) := \sum_{n\in\mathbb{N}} \frac{\varphi(nx_n)}{n}.$$

Observe that $|f((x_n))| \leq \|(x_n)\|_{\ell^1}$. In Exercise 1.4.3, we will show that f is C^1 with differential $df((x_n);(v_n)) = \sum_{n\in\mathbb{N}} v_n \psi(nx_n)$ but not Fréchet differentiable.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

1.28 (Bastiani vs. Fréchet calculus) While Bastiani C^k is weaker than FC^k, Walter (2012, Appendix A.3) shows that there is only a mild loss of differentiability. In particular, $\mathrm{FC}^\infty = C^\infty$. The proofs are somewhat technical induction arguments involving the operator norm. Hence we only summarise the relation between the calculi *on Banach spaces* in the following diagram (arrows denote implications between conditions):

$$C^{k+1} \Longrightarrow \mathrm{FC}^k \underset{\dim E<\infty}{\rightleftarrows} C^k$$

Exercises

1.4.1 Let $(E,\|\cdot\|_E),(F,\|\cdot\|_F)$ be normed spaces and the space $L(E,F)$ of continuous linear maps be endowed with the operator norm $\|A\|_{\mathrm{op}} = \sup_{x\in E\setminus\{0\}} \frac{\|A(x)\|_F}{\|x\|_E}$. Show that the evaluation map $\varepsilon\colon L(E,F)\times E \to F, \varepsilon(A,x) = A(x)$ is continuous.

1.4.2 Let $f\colon U \to F$ be an FC^1-map on $U \subseteq E$, where E,F are Banach spaces.

(a) Show that the Fréchet derivative satisfies $Df(x)(h) = df(x;h)$ for every $x \in U$, $h \in E$ and deduce that f is C^1 in the Bastiani sense.

(b) Use induction to prove that every FC^k-map is already C^k by showing that the kth-Fréchet derivative gives rise to the kth derivative in the Bastiani sense.

1.4.3 We fill in the details for Example 1.27. Notation is as in the example. Prove that

(a) $|\psi(u)| \le 1$ for all $u \in \mathbb{R}$.

(b) $df((x_n);(v_n)) = \sum_{n\in\mathbb{N}} v_n \psi(nx_n)$, hence continuous and thus f is C^1.

(c) $\|df((x_n);\cdot)\|_{\mathrm{op}}$ equals 0 if $x = 0$ but is ≥ 1 if $x_n = 1/n$ for some $n \in \mathbb{N}$.

(d) For $\delta_n := \begin{cases} 0 & \text{if } m \ne n, \\ 1/n & \text{if } m = n, \end{cases}$ the expression $\|df(\delta_n;\cdot)\|_{\mathrm{op}}$ does not converge to 0. Deduce that f is not FC^1.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

1.5 Infinite-Dimensional Manifolds

In this section, we recall the basic notions of manifolds modelled on locally convex spaces. Most of these definitions should be very familiar from the finite-dimensional setting.

1.29 In this section we will write $g \circ f$ as a shorthand for $g \circ f|_{f^{-1}(A)}\colon f^{-1}(A) \to B$ if it helps to avoid clumsy notation.

1.30 Definition (Charts and atlas) Let M be a Hausdorff topological space. A *chart* for M is a homeomorphism $\varphi\colon U_\varphi \to V_\varphi$ from $U_\varphi \subseteq M$ onto $V_\varphi \subseteq E_\varphi$, where E_φ is a locally convex space. Let $r \in \mathbb{N}_0 \cup \{\infty\}$. A *$C^r$-atlas* for M is a set $\mathcal{A}$ of charts for M satisfying the following:

(a) $M = \bigcup_{\varphi\in\mathcal{A}} U_\varphi$.
(b) For all $\varphi, \psi \in \mathcal{A}$ the *change of charts* $\varphi\circ\psi^{-1}$ (which are mappings between open subsets of locally convex spaces) are C^r.[5]

Two C^r-atlases $\mathcal{A}, \mathcal{A}'$ for M are *equivalent* if their union $\mathcal{A} \cup \mathcal{A}'$ is a C^r-atlas for M. This is an equivalence relation.

1.31 Definition A C^r manifold $(M, \mathcal{A})$ is a Hausdorff topological space with an equivalence class of C^r-atlases $\mathcal{A}$. (If the equivalence class $\mathcal{A}$ is clear, we simply write M.)

1.32 Remark In contrast with the finite-dimensional case, we do not require manifolds to be paracompact or second countable (as topological spaces).

In general, the manifolds we are interested in will not be modelled on a single locally convex space. For a C^1-atlas, the locally convex spaces in which charts take their image are necessarily isomorphic on each connected component. However, some examples we will encounter later on have a huge number of connected components. For each of these connected components the locally convex model spaces will, in general, not be isomorphic.

1.33 Example Every locally convex space E is a manifold with global chart given by the identity id_E. Similarly, every $U \subseteq E$ is a manifold with global chart given by the inclusion $U \to E$.

1.34 Example (Hilbert sphere) For a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ the unit sphere $S_H := \{x \in H \mid \langle x, x\rangle = 1\}$ is a C^∞-manifold. To construct charts, we define the Hilbert space $H_{x_0} = \{y \in H \mid \langle y, x_0\rangle = 0\} \subseteq H$ for $x_0 \in S_H$. (If dim

[5] Formally, if the charts do not intersect, the change of charts is the empty map $\emptyset \to \emptyset$ which is C^r.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$H < \infty$, H_{x_0} is a proper subspace of H, but if H is infinite-dimensional, it is isomorphic to H; see Dobrowolski, 1995). Then define the sets $U_{x_0} := \{x \in S_H \mid \langle x, x_0\rangle > 0\}$ and $V_{x_0} := \{y \in H_{x_0} \mid \langle y, y\rangle < 1\}$. We can now define a chart

$$\varphi_{x_0} \colon U_{x_0} \to V_{x_0}, \quad x \mapsto x - \langle x, x_0\rangle x_0$$

(its inverse is given by the formula $\varphi_{x_0}^{-1}(y) = y + \sqrt{1 - \langle y, y\rangle}x_0$). Applying these formulae, we see that the change of charts map for $x_0, z_0 \in S_H$ is a smooth map between open (possibly empty) subsets of Hilbert spaces:

$$\varphi_{z_0} \circ \varphi_{x_0}^{-1}(y) = (y - \langle y, z_0\rangle z_0) + \sqrt{1 - \langle y, y\rangle}(x_0 - \langle x_0, z_0\rangle z_0).$$

1.35 Definition Let M be a C^r-manifold and together with a sequentially closed vector subspace $F_\varphi \subseteq E_\varphi$ for each chart $\varphi \colon U_\varphi \to V_\varphi \subseteq E_\varphi$. A *($C^r$-)submanifold of M* is a subset $N \subseteq M$ such that for each $x \in N$, there exists a chart $\phi \colon U_\phi \to V_\phi$ of M around x such that $\phi(U_\phi \cap N) = V_\phi \cap F_\varphi$. Then $\phi_N := \phi|^{V_\phi \cap F_\varphi}_{U_\phi \cap N}$ is a chart for N, called a *submanifold chart*. Thanks to Lemma 1.25, the submanifold charts form a C^r-atlas for N.

If N is a submanifold of M such that all the sequentially closed subspaces F_ϕ are complemented subspaces of E_ϕ (see §1.7), we call N a *split submanifold* of M.

1.36 Definition Let $(M, \mathcal{A})$ and $(N, \mathcal{B})$ be C^r manifolds. Then the product $M \times N$ becomes a C^r-manifold using the atlas $\mathcal{C} := \{\phi \times \psi \mid \phi \in \mathcal{A}, \psi \in \mathcal{B}\}$. We call the resulting C^r-manifold the *(direct) product of M and N*.

1.37 Definition Let $r \in \mathbb{N}_0 \cup \{\infty\}$ and M, N be C^r-manifolds. A map $f \colon M \to N$ is called C^r if f is continuous and, for every pair of charts ϕ, ψ, the map

$$\psi \circ f \circ \phi^{-1} \colon E_\phi \supseteq \phi(f^{-1}(U_\psi) \cap U_\phi) \to F_\psi$$

is a C^r-map. We write $C^r(M, N)$ for the set of all C^r-maps from M to N.

1.38 Remark Let $f \colon M \to N$ be a continuous map between C^r-manifolds. Assume that for some charts (U_φ, φ) and (U_ψ, ψ) the composition $\varphi \circ f \circ \psi$ is $C^r, r \in \mathbb{N} \cup \{\infty\}$. Then for any other pair of charts (U_κ, κ) and (U_λ, λ) with $f(U_\lambda \cap U_\psi) \subseteq U_\varphi \cap U_\kappa$ we have on $U_\lambda \cap U_\psi$ that

$$\kappa \circ f \circ \lambda^{-1}|_{\lambda(U_\lambda \cap U_\psi)} = (\kappa \circ \varphi^{-1}) \circ (\varphi \circ f \circ \psi^{-1}) \circ (\psi \circ \lambda^{-1}),$$

where the mapping in the middle is C^r by assumption and the other mappings are change of charts (which are C^r by M, N being C^r-manifolds). Hence $\kappa \circ f \circ \lambda^{-1}|_{\lambda(U_\lambda \cap U_\psi)}$ is also C^r by the chain rule, Proposition 1.23. This argument is called 'insertion of charts' and we leave it from now on to the reader. With the insertion of charts argument, it is easy to see that:

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(a) It suffices to test the C^r-property with respect to any atlas of M and N.
(b) If $f\colon M \to N$ and $g\colon N \to L$ are C^r-maps, so is $g \circ f\colon M \to L$.
(c) Using Lemma 1.24: If M, N_1, N_2 are C^r-manifolds and $f_i\colon M \to N_i$, $i = 1,2$ are mappings. Then $f \coloneqq (f_1, f_2)\colon M \to N_1 \times N_2$ is C^r if and only if f_1, f_2 are C^r.

1.39 Lemma *Let N be a (C^r-)submanifold of the C^r-manifold M. Then the inclusion $\iota\colon N \to M$ is C^r. Further, $f\colon P \to N$ is C^r if and only if $\iota \circ f$ is C^r.*

Proof Thanks to Remark 1.38, it suffices to check the C^r-property of ι in charts that cover N. Thus we may choose charts (ϕ, U_ϕ) of M that induce submanifold charts ϕ_N as in Definition 1.35. But then $\phi \circ \iota \circ \phi_N^{-1}$ is the inclusion map $V_\phi\colon F \to V$ which is C^r (as restriction of a continuous linear map).

If f is C^r, so is $\iota \circ f$ by Remark 1.38. Conversely, let $\iota \circ f$ be a C^r-map and ϕ, ϕ_N as before and ψ a chart for P. Then $\phi \circ \iota \circ f \circ \psi^{-1}\colon \psi(U_\psi \cap f^{-1}(U_\phi)) \to E$ is C^r with values in the sequentially closed subspace F and thus $(\phi \circ \iota \circ f \circ \psi)|^F = \phi_N \circ f \circ \psi^{-1}$ is C^r by Lemma 1.25. We conclude that f is C^r. □

Exercises

1.5.1 Let $f\colon M \to N$ be a C^r-map between C^r-manifolds. Show that the $\operatorname{graph}(f) \coloneqq \{(m, f(m)) \mid m \in M\}$ is a split submanifold of $M \times N$.
Hint: Use the description of the graph to construct submanifold charts by hand.

1.5.2 Verify the details of Example 1.34: Check that the charts make sense as mappings from U_{x_0} to V_{x_0}. Show that the change of charts $\varphi_{x_0} \circ \varphi_{z_0}^{-1}$ is smooth for all $x_0, z_0 \in S_H$ such that $U_{x_0} \cap U_{z_0} \neq \emptyset$.

1.5.3 Let M be a manifold and $U \subseteq M$. Let $\mathcal{A}$ be an atlas of M. Endow U with the subspace topology and show that $\mathcal{A}_U \coloneqq \{\phi|_{U \cap U_\phi} \mid (\phi, U_\phi) \in \mathcal{A}\}$ is a manifold atlas for U turning it into a submanifold of M.

1.5.4 Check that the set C in Definition 1.36 is a C^r-atlas for the product manifold.

1.5.5 Show that a compact manifold must be modelled on a finite-dimensional space.
Hint: Proposition A.3.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

1.6 Tangent Spaces and the Tangent Bundle

1.40 Definition Let M be a C^r-manifold ($r \geq 1$) and $p \in M$. We say that a C^1-curve γ *passes through* p if $\gamma(0) = p$. For two such curves γ, η we define the relation

$$\gamma \sim \eta \Leftrightarrow (\phi \circ \gamma)'(0) = (\phi \circ \eta)'(0) \tag{1.8}$$

for some chart ϕ of M around p. By the chain rule, (1.8) holds for every chart around p and defines an equivalence relation on the set of all curves passing through p. The equivalence class $[\gamma]$ is called the *(geometric) tangent vector* (of M at p). Define the *(geometric) tangent space of M at p* as the set T_pM of all geometric tangent vectors at p.

We recall now that the tangent space at a point can be turned into a locally convex space isomorphic to the modelling space at that point.

1.41 Lemma (a) *Let ϕ be a chart of M around p, set $p_\phi := \phi(p)$. Then*

$$h_\phi : E_\phi \to T_pM, \quad h_\phi(y) := [t \mapsto \phi^{-1}(p_\phi + ty)]$$

is a bijection with inverse $h_\phi^{-1} : T_pM \to E_\phi$, $[\gamma] \mapsto (\phi \circ \gamma)'(0)$.

(b) *For all charts ϕ, ψ around p, we have $h_\psi^{-1} \circ h_\phi = d(\psi \circ \phi^{-1})(p_\phi; \cdot)$ which is an automorphism of the topological vector space E.*

(c) *T_pM admits a unique locally convex space structure such that h_ϕ is an isomorphism of locally convex spaces for some (and hence all) charts ϕ of M around p.*

Proof (a) Note that h_ϕ and h_ϕ^{-1} are well defined and h_ϕ^{-1} is injective. For $y \in E_\phi$, $h_\phi^{-1} \circ h_\phi(y)) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \phi(\phi^{-1}(p_\phi + ty)) = y$. Thus h_ϕ^{-1} is surjective and the inverse of h_ϕ.

(b) Compute for $y \in E$: $h_\psi \circ h_\phi^{-1}(y) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \psi(\phi^{-1}(p_\phi + ty)) = d(\psi \circ \phi^{-1})(p_\phi; \cdot)$.

(c) This follows directly from the definition of the vector space structure. □

1.42 Let $U \subseteq E$, E a locally convex space and $f : U \to F$ a C^1-map. We define the mapping

$$Tf : U \times E \to F \times F, \quad (x, v) \mapsto (f(x), df(x; v))$$

and call this mapping the *tangent map* of f. Note that the chain rule, Proposition 1.23, can now be written as $T(f \circ g) = Tf \circ Tg$.

1.43 Definition (Tangent bundle) Let $(M, \mathcal{A})$ be a C^r-manifold with $r \geq 1$. We call $TM := \bigcup_{p \in M} T_pM$ the *tangent bundle* of M. Then $\pi_M : TM \to M$,

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$T_pM \ni v \mapsto p$ is called the *bundle projection*. We equip TM with the final topology with respect to the family $(T\phi^{-1})_{\phi\in\mathcal{A}}$ of mappings

$$T\phi^{-1}\colon V_\phi \times E_\phi \to TM, \quad (x,v) \mapsto [t \mapsto \phi^{-1}(x+ty)] \in T_{\phi^{-1}(x)}M.$$

Note that $TU_\phi = \pi_M^{-1}(U_\phi)$ is open in TU for all $\phi \in \mathcal{A}$, and

$$T\phi := (T\phi^{-1})^{-1}\colon TU_\phi \to V_\phi \times E$$

is a homeomorphism. Moreover, $\mathcal{B} := \{T\phi \mid \phi \in \mathcal{A}\}$ is a C^{r-1}-atlas for TM. Thus TM becomes a C^{r-1}-manifold and $\pi_M\colon TM \to M$ a C^{r-1}-map.

1.44 Lemma *We check the details in Definition 1.43. Let $\varphi,\psi \in \mathcal{A}$:*

(a) *We have $T(\psi\circ\phi^{-1})\circ T\phi = T\psi$.*
(b) *TU_φ is open in TM and $T\varphi\colon TU_\varphi \to V_\varphi \times E_\varphi$ is a homeomorphism.*
(c) *$\mathcal{B} = \{T\phi \mid \phi \in \mathcal{A}\}$ is a C^{r-1}-atlas, if M is Hausdorff, so is TM.*
(d) *π_M is a C^{r-1}-map.*
(e) *TM induces on each tangent space T_pM its natural topology.*

Proof (a) This follows from Lemma 1.41(b); we leave the details to the reader.

(b) If $\varphi,\psi \in \mathcal{A}$, we have $(T\psi^{-1})^{-1}(TU\varphi) = T\psi(TU_\varphi \cap TU_\psi) = \psi(U_\varphi\cap U_\psi)\times E_\psi \subseteq V_\psi \times E_\psi$. By the definition of the final topology, TU_φ is open in TM.

By definition of the final topology $T\varphi^{-1}$ is continuous, and so $T\varphi$ is open for every $\varphi \in \mathcal{A}$. For continuity, pick $U \subseteq V_\varphi \times E_\varphi$ and let $\psi \in \mathcal{A}$. Now $W := U \cap \varphi(U_\varphi\cap U_\psi)\times E_\varphi \subseteq V_\varphi \times E_\varphi$, whence a quick computation shows

$$(T\psi^{-1})^{-1}((T\varphi)^{-1}(U)) = T(\psi\circ\varphi^{-1})(W) \subseteq V_\psi \times E_\psi$$

as $T(\psi\circ\varphi^{-1})$ is a homeomorphism between open subsets of $V_\varphi \times E_\varphi$ and $V_\psi \times E_\psi$. We deduce that $(T\varphi)^{-1}(U)$ is open in TM, and so $T\varphi$ is continuous.

(c) By (b) each $T\phi$ is a homeomorphism from an open subset of TM onto an open subset of $E_\phi \times E_\phi$, whence it is a chart. Clearly, the TU_φ cover TM and by (a) the transition maps are $T(\psi\circ\varphi^{-1}) = (\psi\circ\varphi^{-1}, d(\psi\circ\varphi^{-1}))$ whence C^{r-1}. The Hausdorff property is left as Exercise 1.6.1.

(d) In every chart we have $\varphi\circ\pi_M = \mathrm{pr}_1\circ T\varphi$, where $\mathrm{pr}_1\colon V_\varphi \times E_\varphi \to V_\varphi$ is the canonical projection. As the charts conjugate π_M to a smooth map, it is of class C^{r-1}.

(e) By the definition of the vector topology in Lemma 1.41 (c) this follows from (b). □

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

1.45 Remark In finite-dimensional differential geometry (and on Banach manifolds), one often introduces the *dual bundle* or *cotangent bundle*

$$T^*M := L(TM,\mathbb{R}) := \bigcup_{x\in M}(T_xM)' = \bigcup_{x\in M} L(T_xM,\mathbb{R}),$$

where $(T_xM)' := L(T_xM,\mathbb{R})$ is the space of continuous linear mappings $T_xM \to \mathbb{R}$ (with the operator norm topology if T_xM is Banach; see Klingenberg, 1995 or Lang, 1999, III.1 for the bundle structure). In our setting, there is, *in general*, no canonical manifold structure which turns T^*M into a vector bundle.

This poses a problem for theories employing the dual bundle, for example, symplectic geometry, and also differential forms cannot be defined as sections of the dual bundle. However, in the case of differential forms, one can circumvent this problem and obtain a theory similar to finite-dimensional differential forms without the dual bundle. We briefly discuss this in Appendix E.

We will now introduce tangent mappings of differentiable mappings between manifolds (see Definition 1.42 for the case of an open subset of a locally convex space).

1.46 Definition (Tangent maps) Let $f\colon M \to N$ be a C^r-map between C^r-manifolds for $r \geq 1$. Define the mappings

$$T_pf\colon T_pM \to T_{f(p)}N, \quad [\gamma] \mapsto [f\circ\gamma], \quad p \in M.$$

Then we define the *tangent map* $Tf\colon TM \to TN$, $T_pM \ni [\gamma] \mapsto T_pf([\gamma])$. Note that by construction $\pi_N \circ Tf = f \circ \pi_M$. Moreover, for each pair of charts ψ of N and ϕ of M such that $f(U_\phi) \subseteq U_\psi$, the following diagram is commutative:

$$\begin{array}{ccc} TU_\phi & \xrightarrow{Tf|_{TU_\phi}^{TU_\psi}} & TU_\psi \\ \downarrow{\scriptstyle T\phi} & & \downarrow{\scriptstyle T\psi} \\ V_\phi \times E_\phi & \xrightarrow{T(\psi\circ f\circ\phi^{-1})} & V_\psi \times F_\psi. \end{array}$$

Hence the tangent map Tf is a C^{r-1}-map if f is a C^r-map.

1.47 Lemma (Chain rule on manifolds) *Let M,N,L be C^r-manifolds and $f\colon M \to N$, $g\colon N \to L$ be C^r-maps with $r \geq 1$. Then $T(g\circ f) = Tg \circ Tf$.*

Note that we can of course iterate the tangent construction and form the higher tangent manifolds $T^kM := \underbrace{(T(T(\cdots(T}_{k \text{ times}}M)\cdots)$ if M is a C^ℓ-manifold and $k \leq \ell$. Similarly one defines higher tangent maps $T^kf := \underbrace{(T(T(\cdots(T}_{k \text{ times}}f)\cdots)$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

For later use, we set a notation for derivatives of manifold valued curves.

1.48 Definition Let M be a manifold and $c\colon J \to M$ be a C^1-map from some interval $J \subseteq \mathbb{R}$. Then we identify $T_t J = \mathbb{R}$ and define the mapping

$$\dot{c}\colon J \to TM, \quad t \mapsto Tc(t,1).$$

Note that $\dot{c}$ is the manifold version of the curve differential c', in particular, if (U,φ) is a chart of M, we have for each $t \in c^{-1}(U)$ the relation $T\varphi(\dot{c}(t)) = (\varphi \circ c, (\varphi \circ c)')$.

Exercises

1.6.1 Verify the details in Definition 1.46 and supply a proof for Lemma 1.47. Show, in particular, that

(a) Tf is well defined and defines a C^{r-1}-map with the claimed properties.

(b) Check that for $M = U$ an open subset of a locally convex space, both definitions of tangent mappings coincide.

(c) The manifold TM is a Hausdorff topological space.
Hint: Consider two cases for $v, w \in TM$: $\pi_M(v) = \pi_M(w)$ and $\pi_M(v) \neq \pi_M(w)$.

1.6.2 Show inductively, that

(a) If E is a locally convex space and $U \subseteq E$, then $T^k U \cong U \times E^{2^k-1}$.

(b) For $U \subseteq E, V \subseteq F$ in locally convex spaces, we have

$$d^k f(x; v_1, \ldots, v_k) = \mathrm{pr}_{2^k}\left(T^k f(x, w_1, \ldots, w_{2^k-1}\right),$$

where pr_{2^k} is the projection onto the 2^kth component and $w_{2^i+1} = v_{i+1}$ for $0 \leq i \leq k-1$ and $w_i = 0$ else.

1.6.3 Establish a manifold version of the rule on partial differentials Proposition 1.20, that is, show that: If $f\colon M_1 \times M_2 \to N$ is a C^1-map (between C^1-manifolds) and $p_i\colon M_1 \times M_2 \to M_i$ are the canonical projections, then for $p = (x,y) \in M_1 \times M_2$,

$$T_p f(v) = T_x f(\cdot, y)(Tp_1(v)) + T_y f(x, \cdot)(Tp_2(v)).$$

Identifying $T(M_1 \times M_2)$ with $TM_1 \times TM_2$ and $v = (v_x, v_y)$, this formula becomes

$$T_p f = T_x f(\cdot, y)(v_x) + T_y f(x, \cdot)(v_y).$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

1.7 Elements of Differential Geometry: Submersions and Immersions

Immersions and submersions are among the first tools students encounter in courses on differential geometry when asked to construct (sub-)manifolds. They can still serve this purpose in infinite dimensions if one chooses one's definitions carefully. In this section, we follow Hamilton (1982, Section 4.4)[6] and the discussion requires the concept of complemented subspaces of a locally convex space (see Appendix A.3).

1.49 Definition Let M, N be smooth manifolds and $\phi\colon M \to N$ smooth. We say that ϕ is

(a) an *immersion* if for every $x \in M$ there are manifold charts κ, ψ around x and $\phi(x)$ such that $F_\psi \cong E_\kappa \times H$ (as locally convex spaces; see Appendix A.3), the local representative of ϕ in these charts is the inclusion $E_\kappa \to E_\kappa \times H \cong F_\psi$,
(b) an *embedding* if f is an immersion and a topological embedding (i.e. a homeomorphism onto its image),
(c) a *submersion* if for every $x \in M$ there are manifold charts κ, ψ around x and $\phi(x)$ such that $E_\kappa \cong F_\psi \times H$ and the local representative of ϕ in these charts is the projection $E_\kappa \cong F_\psi \times H \to F_\psi$.

1.50 Lemma *If $f\colon M \to N$ is an immersion, then for every $x \in M$ there is an open neighbourhood W_x such that $f|_{W_x}$ is an embedding.*

Proof Pick immersion charts around x and $f(x)$, that is, charts $(U, \varphi), x \in U$ and (W, ψ) such that $f(x) \in W$ and $\psi \circ f \circ \varphi^{-1} = j$, where $j\colon E_\varphi \to E_\psi \cong E_\varphi \times F$ is the inclusion of the complemented subspace E_φ of E_ψ. Note that j is a topological embedding onto its image, whence $f|_{U_\varphi} = \psi^{-1} \circ j \circ \varphi$ is a topological embedding (and thus an embedding). □

There are several alternative characterisations of submersions and immersions known in the finite-dimensional setting. Many of these turn out to be weaker in our setting.

1.51 Lemma *A smooth map $f\colon M \to N$ is a submersion if and only if for each $x \in M$ there are (U, φ) of M and (V, ψ) of N with $x \in U$ and $f(U) \subseteq V$ such that $\psi \circ f \circ \varphi^{-1} = \pi|_{\varphi(U)}$ for a continuous linear map π with continuous linear right inverse σ (i.e. $\pi \circ \sigma = \mathrm{id}$).*

[6] The smoothness assumption conveniently shortens the exposition but can of course be replaced by finite orders of differentiability; see Glöckner (2016).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Proof If f is a submersion and $x \in M$, we can pick submersion charts around x, that is, charts $\varphi\colon U_\varphi \to U_F \times U_H \subseteq F \times H \cong E$ with $U_F \subseteq F$ and $U_H \subseteq E$ and $\psi\colon V_\psi \to U_F$ such that $\psi \circ f \circ \varphi^{-1}(x,y) = x$, for all $(x,y) \in U_F \times U_H$. Obviously, the projection is continuous linear with right inverse given by the inclusion $F \to F \times H \cong E$.

Let us conversely assume that around $x \in M$ there are charts such that $\psi \circ f \circ \varphi^{-1} = \pi|_{\varphi(U)}$ holds for a continuous linear map $\pi\colon E \to F$ with continuous right inverse $\sigma\colon F \to E$. Then $\pi|_{\sigma(F)}\sigma(F) \to F$ is an isomorphism of locally convex spaces with inverse σ. We obtain a new chart $\nu := \sigma \circ \psi$ of N. Now $\kappa := \sigma \circ \pi\colon E \to E$ is continuous linear and satisfies $\kappa \circ \kappa = \kappa$. We deduce from Lemma A.21 that $\sigma(F)$ is a complemented subspace of E and can identify $E \cong \sigma(F) \times \ker(\pi)$ such that κ becomes the projection onto $\sigma(F)$. Shrinking U and V if necessary we may assume $\varphi(U) = A \times B$ and $\nu(V) = A$ for open sets $A \subseteq \sigma(F)$ and $B \subseteq \ker\pi$. Then

$$\nu \circ f \circ \varphi^{-1} = \sigma \circ \psi \circ f \circ \varphi^{-1} = \sigma \circ \pi|_{A\times B} = \kappa|_{A\times B}. \qquad \square$$

1.52 Remark The alternative characterisation of submersions from Lemma 1.51 implies (see Exercise 1.7.5) that a submersion admits local smooth sections. If the manifolds M and N are Banach manifolds (i.e. modelled on Banach spaces), the inverse function theorem implies that the existence of local smooth sections is even equivalent to being a submersion (see Margalef-Roig and Domínguez, 1992, Proposition 4.1.13).

Finally, we mention that *if the submersion φ is surjective*, we obtain the following result: *a map $f\colon N \to L$ is C^r if and only if $f \circ \varphi$ is a C^r-map* for $r \in \mathbb{N}_0 \cup \{\infty\}$ (see Exercise 1.7.6).

In finite-dimensional differential geometry, the above definitions are usually not the definitions of submersions/immersions but one deduces them from 'easier conditions' involving the tangent mappings, such as the following.

1.53 Definition Let M, N be smooth manifolds and $\phi\colon M \to N$ smooth. We say that ϕ is

(a) *infinitesimally injective (or surjective)* if the tangent map $T_x\phi\colon T_xM \to T_{\phi(x)}N$ is injective (or surjective, respectively) for every $x \in M$,
(b) a *naïve immersion* if for every $x \in M$ the tangent map $T_x\phi\colon T_xM \to T_{\phi(x)}N$ is a topological embedding onto a complemented subspace of $T_{\phi(x)}N$,
(c) a *naïve submersion* if for every $x \in M$ the map $T_x\phi\colon T_xM \to T_{\phi(x)}N$ has a continuous linear right inverse.

In infinite dimensions, none of the naïve, infinitesimal versions or the properties from Definition 1.49 are equivalent as the following (counter-)examples show.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

1.54 (Infinitesimal properties are weaker than naïve properties) Consider the Banach space c_0 of all (real) sequences converging to 0 as a subspace of the Banach space ℓ^∞ of all bounded real sequences. We obtain a *short exact sequence*[7] of Banach spaces

$$0 \longrightarrow c_0 \xhookrightarrow{i} \ell^\infty \xrightarrow{q} \ell^\infty/c_0 \longrightarrow 0. \tag{1.9}$$

As i and q are continuous linear we have $T_x i(y) = di(x,y) = i(y)$ and $T_x q(y) = q(y)$, whence i is infinitesimally injective and q is infinitesimally surjective. Since c_0 is not complemented, Example A.23, i is not a naïve immersion. This implies that (1.9) does not split, that is, q does not admit a continuous linear right inverse and thus cannot be a naïve submersion.

A submersion as in Definition 1.49 turns out to be an open map (Glöckner, 2016, Lemma 1.7). In finite-dimensional differential geometry, this is a consequence of the inverse function theorem (when applied to a naïve submersion). Going beyond Banach manifolds, the submersion property is stronger than the naïve notion.

1.55 (Naïve submersions need not be submersions) Consider the space $A := C(\mathbb{R},\mathbb{R})$ of continuous functions from the reals to the reals. The pointwise operations and the compact open topology (see Appendix B.2) turn A into a locally convex space (by Lemma B.7, one can even show that it is a Fréchet space). Then the map $\exp_A\colon A \to A$, $f \mapsto e^f$ is continuous by Lemma B.8. Recall that the point evaluations $\mathrm{ev}_x(f) := f(x)$ are continuous linear on A, whence we can use them to find a candidate for the derivative of $\exp_A$. Then the finite-dimensional chain rule yields the only candidate for the derivative of $\exp_A$ to be $d\exp_A(f;g)(x) = g(x)\cdot\exp_A(f)(x)$. Computing in the seminorms, one can show that indeed $d\exp_A(f;g) = g\cdot\exp_A(f)$ and by induction $\exp_A$ is smooth. Moreover, at the constant zero function $\mathbf{0}$ we have $d\exp_A(\mathbf{0};\cdot) = \mathrm{id}_A$, so $\exp_A$ is a naïve submersion. However, $\exp_A$ takes values in $C(\mathbb{R},]0,\infty[)$ which do not contain a neighbourhood of $\exp_A(\mathbf{0}) = \mathbf{1}$. Thus $\exp_A$ cannot be a submersion (as all submersions are open mappings by Glöckner (2016, Lemma 1.7).

[7] In the category of locally convex spaces, a sequence

$$0 \longrightarrow A \xrightarrow{i} B \xrightarrow{q} C \longrightarrow 0$$

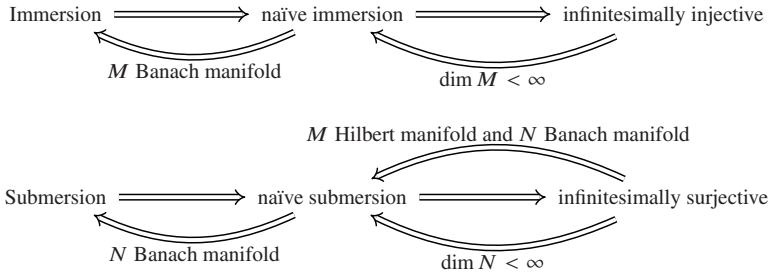
of continuous linear maps is *exact* if it satisfies both of the following conditions:
(a) *algebraically exact*, that is, images of maps coincide with kernels of the next map;
(b) *topologically exact*, that is, i and q are open mappings onto their images.
If A, B and C are Fréchet (or Banach) spaces, topological exactness follows from algebraic exactness by virtue of the open mapping theorem (Rudin, 1991, I. 2.11); for general locally convex spaces this is not the case.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

This example also shows that the Inverse Function Theorem fails in this case (Eells, 1966); see also Appendix A.5.

However, as Glöckner (2016) shows, the following relations do hold.

1.56 (Submersions, immersions vs. the naïve and infinitesimal concepts) Consider a map $\phi\colon M \to N$ between manifolds modelled on locally convex spaces. Then:



The reason one really would like the strong notions of submersions and immersions is that these notions are strong enough to carry over the usual statements on submersions and immersions to the setting of infinite-dimensional manifolds. For example, one can prove several useful statements on split submanifolds. Again we refer to Glöckner (2016) for more general results on submersions and immersions in infinite dimensions.

1.57 Definition Let $f\colon M \to N$ be smooth and $S \subseteq N$ be a split submanifold. Then f is *transversal over* S if for each $m \in f^{-1}(S)$ and submanifold chart $\psi\colon V \to V_1 \times V_2$ with $\psi(f(m)) = (0,0)$ and $\psi(S) \subseteq V_1 \times \{0\}$, there exists an open m-neighbourhood U with $f(U) \subseteq V$ and

$$U \xrightarrow{f} V \xrightarrow{\psi} V_1 \times V_2 \xrightarrow{\mathrm{pr}_2} V_2 \tag{1.10}$$

is a submersion.

Note that due to the fact that compositions of submersions are again submersions (see Exercise 1.7.1), if f is a submersion, (1.10) is always a submersion.

1.58 Proposition *Let $\phi\colon M \to N$ be a smooth map. If $S \subseteq N$ is a split submanifold*[8] *of N such that f is transversal over S, then $\phi^{-1}(S)$ is a submanifold of M.*

[8] A more involved proof works for every submanifold (not only for split ones), see Glöckner (2016, Theorem C).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Proof The map (1.10) is a submersion. Shrinking U, V_2, there are charts $\varphi\colon U \to U_1 \times U_2$ and $\kappa\colon V_2 \to U_2$ such that $\varphi(m) = (0,0)$ and $\kappa(0) = 0$ and the following commutes:

$$\begin{array}{ccccccc} U & \xrightarrow{f} & V & \xrightarrow{\psi} & V_1 \times V_2 & \xrightarrow{\mathrm{pr}_2} & V_2 \\ \downarrow{\scriptstyle\varphi} & & & & & & \downarrow{\scriptstyle\kappa} \\ U_1 \times U_2 & & & & \xrightarrow{\qquad\mathrm{pr}_2\qquad} & & U_2. \end{array}$$

Now we will prove that φ is a submanifold chart for $f^{-1}(S)$, that is, $\varphi(U \cap f^{-1}(S)) = \varphi(U) \cap (U_1 \times \{0\})$. To see this, note that since ψ is a submanifold chart, we have for $x \in U$ that $f(x) \in S$ if and only if $\mathrm{pr}_2(\psi(f(x))) = 0$. Now the commutativity of the diagram shows that this is the case if and only if $\varphi(x) \in \mathrm{pr}_2^{-1}(0) = U_1 \times \{0\}$. □

1.59 Corollary *If $f\colon M \to N$ is a submersion, $f^{-1}(n)$ is a split submanifold for $n \in N$.*

1.60 Lemma *Let $f\colon M \to P$ and $g\colon N \to P$ be smooth maps and g be a submersion. Then the* fibre product $M \times_P N := \{(m,n) \in M \times N \mid f(m) = g(p)\}$ *is a split submanifold of $M \times N$ and the projection $\mathrm{pr}_1\colon M \times_P N \to M$ is a submersion.*

Proof Let $(m,n) \in M \times_P N$ and pick submersion charts $\psi\colon U_\psi \to V_\psi \subseteq E_N$, $\kappa\colon U_\kappa \to V_\kappa \subseteq E_P$ for g with $n \in U_\psi$. Recall that for the submersion charts $\kappa \circ g \circ \psi^{-1} = \pi$ for a continuous projection $\pi\colon E_N = E_P \times F \to E_P$ (where F is the subspace complement) and we may assume that $V_\psi = V_\kappa \times \tilde{V}$. Finally, we pick a chart (U_φ, φ) of M such that $m \in U_\varphi$ and $f(U_\varphi) \subseteq U_\kappa$. Hence we obtain a commutative diagram

$$\begin{array}{ccccc} & & U_\psi & \xrightarrow{\psi} & V_\psi = V_\kappa \times \tilde{V} \\ & & \downarrow{\scriptstyle g} & \searrow{\scriptstyle\pi} & \downarrow{\scriptstyle\mathrm{pr}_{V_\kappa}} \\ U_\varphi & \xrightarrow{f} & U_\kappa & \xrightarrow{\kappa} & V_\kappa. \end{array}$$

Denote by $\mathrm{pr}_{\tilde{V}}$ the projection onto $\tilde{V}$. Then we construct a smooth map for $(m,n) \in U_\varphi \times U_\psi$ via

$$\delta(m,n) := (\varphi(m), (\mathrm{pr}_{V_\kappa}(\psi(n)) - \kappa(f(m)), \mathrm{pr}_{\tilde{V}}(\psi(n)))) \in V_\varphi \times ((V_\kappa - V_\kappa) \times \tilde{V}).$$

This mapping is smoothly invertible with inverse given by

$$\delta^{-1}(x,(y,z)) = (\varphi^{-1}(x), \psi^{-1}(y + \kappa(f(\varphi^{-1}(x))), z)).$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

We leave it again as an exercise to work out that the domain of the inverse is open. Note that due to the commutative diagram we see that $(m,n) \in M \times_P N$ if and only if (m,n) maps under δ to $E_M \times \{0\} \times F$ which is a complemented subspace of $E_M \times E_P \times F \cong E_M \times E_N$. Thus $M \times_P N$ is a split submanifold of $M \times N$. Since the projection $\mathrm{pr}_1 \colon M \times N \to N$ is smooth, so is its restriction to $M \times_P N$ by Lemma 1.39. As $\mathrm{pr}_1 \colon M \times_P N \to M$ is conjugated by δ and φ to the projection pr_{V_κ}, it is a submersion. □

Finally, there is a close connection between embeddings and split submanifolds.

1.61 Lemma *Let $f \colon M \to N$ be smooth. The following conditions are equivalent:*

(a) *f is an embedding.*
(b) *$f(M)$ is a split submanifold of N and $f|^{f(M)} \colon M \to f(M)$ is a diffeomorphism.*

Proof Let E, F be the modelling spaces of M and N, respectively.

(a) $\Rightarrow$ (b): By assumption, f is a topological embedding and a smooth immersion. Consider $y \in f(M)$ and $x \in M$ with $f(x) = y$ and pick charts $\varphi_x \colon U_x \to V_x \subseteq E$ and $\varphi_y \colon U_y \to V_y \subseteq F$ such that $x \in U_x$, $f(U_x) \subseteq U_y$ and $\varphi_y \circ f \circ \varphi_x^{-1} = j|_{V_x}$ for a linear topological embedding $j \colon E \to F$ onto a complemented subspace $j(E) \times H = F$. Since $j(V_x)$ is relatively open, we may adjust choices such that $j(V_x) = V_y \cap j(E)$. A quick computation then shows that φ_y restricts to a submanifold chart and $f(M)$ becomes a split submanifold of N. Moreover, in the (sub)manifold charts we have $j|_{V_x}^{V_y \cap j(E)} = \varphi_y|_{f(M) \cap U_y} \circ f \circ \varphi_x^{-1}$ and this map is a diffeomorphism. Thus $f|^{f(M)} \colon M \to f(M)$ is a local diffeomorphism and a homeomorphism, hence a diffeomorphism.

(b) $\Rightarrow$ (a) Let $\iota \colon f(M) \to N$ be the inclusion map. Since $f|^{f(M)}$ is a diffeomorphism, $\iota \circ f|^{f(M)}$ is a topological embedding. Since $f(M)$ is a split submanifold, there is an isomorphism $\alpha \colon E \to \alpha(E) \subseteq F$ of locally convex spaces such that $\alpha(E)$ is complemented in F. Pick charts $\varphi_x \colon U_x \to V_x$ and $\varphi_{f(x)} \colon U_{f(x)} \to V_{f(x)}$ with $x \in U_x$ and $f(U_x) \subseteq U_{f(x)}$. We may assume that $V_{f(x)} = P \times Q$ for $P \Subset \alpha(E)$ and $\varphi_{f(x)}(U_x \cap f(M)) = V_{f(x)} \cap \alpha(E) = P$. Set now $W := \alpha^{-1}(P)$. Then it is easy to see that $\theta := (\varphi_{f(x)} \circ f \circ \varphi_x^{-1})^{-1} \circ \alpha|_W \colon W \to V_x$ makes sense and is a smooth diffeomorphism, whence $\theta^{-1} \circ \varphi_x \colon U_x \to W$ is a chart for M. By construction, $\varphi_{f(x)} \circ f \circ (\theta^{-1} \circ \varphi_x)^{-1} = \alpha|_W$ is a linear topological embedding onto $\alpha(E)$. This shows that f is an immersion. □

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

1.7.1 Exercises

1.7.1 Show that if $f\colon M \to N$ and $g\colon N \to L$ are submersions, so is $g \circ f\colon M \to L$.
Hint: The submersion property is local. Try constructing small enough neighbourhoods around $m \in U \subseteq M, f(m) \in V \subseteq N$ and $g(f(m)) \in W \subseteq L$ and charts such that:

$$\begin{array}{ccccc} U & \xrightarrow{\quad f \quad} & V & \xrightarrow{\quad g \quad} & W \\ \downarrow & & \downarrow & & \downarrow \\ F \times X & \xrightarrow{\mathrm{pr}_F} & F \cong Y \times Z & \xrightarrow{\mathrm{pr}_Y} & Y. \end{array}$$

1.7.2 Let $f\colon M \to N$ and $g\colon K \to L$ be submersions (immersions). Show that then $f \times g\colon M \times K \to N \times L$, $(m,k) \mapsto (f(m), g(k))$ is also a submersion (immersion).

1.7.3 Let $p\colon M \to N$ be a submersion and $n \in N$. From Corollary 1.59 we obtain a submanifold $P := p^{-1}(n)$. Show that for $x \in P$ one can identify the tangent space of the submanifold as $T_x P = \ker T_x p = \{v \in T_x M \mid T_x p(v) = 0\}$.

1.7.4 Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Prove that $\psi\colon H \setminus \{0\} \to \mathbb{R}$, $x \mapsto \langle x, x\rangle$ is a submersion and deduce that the Hilbert sphere $S_H = \psi^{-1}(1)$ is a submanifold of H. Then show that $T_x S_H = \{v \in H \mid \langle v, x\rangle = 0\}$ for all $x \in S_H$ (that is, the tangent space is the orthogonal complement to the base point x).

1.7.5 Let $\varphi\colon M \to N$ be a smooth submersion. Show that φ admits smooth local sections, that is, for every $x \in M$ there is $\varphi(x) \in U \subseteq N$ and a smooth map $\sigma\colon U \to M$ with $\sigma(\varphi(x)) = x$ and $\varphi \circ \sigma = \mathrm{id}_U$. Deduce then that φ is an open map.
Hint: Use the characterisation from Remark 1.52.
Remark: If M, N are Banach manifolds, the existence of local sections is equivalent to φ being a submersion; see Margalef-Roig and Domínguez (1992, Proposition 4.1.13).

1.7.6 Let $\varphi\colon M \to N$ be a smooth surjective submersion. Show that $f\colon N \to L$ is C^r if and only if $f \circ \varphi$ is C^r for $r \in \mathbb{N}_0 \cup \{\infty\}$.
Hint: Use Exercise 1.7.5.

1.7.7 Show that if $f\colon M \to N$ and $g\colon N \to L$ are immersions (embeddings), so is $g \circ f$.

1.7.8 Work out the details omitted in the proof of Lemma 1.61.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

2

Spaces and Manifolds of Smooth Maps

In this chapter, we consider spaces of differentiable mappings as infinite-dimensional spaces. These spaces will then serve as the model spaces for manifolds of mappings, that is, manifolds of differentiable mappings between manifolds.

2.1 Topological Structure of Spaces of Differentiable Mappings

In this section, we denote by M, N (possibly infinite-dimensional) manifolds.

2.1 Definition Endow the space $C^\infty(M,N)$ with the initial topology with respect to the map

$$\Phi\colon C^\infty(M,N) \to \prod_{k\in\mathbb{N}_0} C(T^kM, T^kN)_{\text{c.o.}}, \quad f \mapsto (T^k f)_{k\in\mathbb{N}_0},$$

where the spaces on the right-hand side carry the compact open topology (see §B.2). The resulting topology on $C^\infty(M,N)$ is called the *compact open C^∞-topology*.

2.2 Remark (a) The map Φ is clearly injective (as $T^0 f := f$). Therefore, Φ is a homeomorphism onto its image.

(b) Note that the compact open C^∞-topology is also the initial topology with respect to the mappings

$$T^k\colon C^\infty(M,N) \to C(T^kM, T^kN)_{\text{c.o}}, \quad f \mapsto T^k f, \quad k \in \mathbb{N}_0.$$

(c) If $M \subseteq E, N \subseteq F$ for E, F locally convex spaces, the compact open C^∞-topology is the initial topology with respect to the mappings

$$d^k\colon C^\infty(M,N) \to C(M\times E^k, N)_{\text{c.o.}}, \quad f \mapsto d^k f, \quad k \in N_0.$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(d) By construction the compact open C^∞-topology is finer than the compact open topology (i.e. the topology induced by the inclusion $C^\infty(M,N) \to C(M,N)$). In particular, if M is locally compact (i.e. M is finite dimensional), the evaluation $\mathrm{ev}\colon C^\infty(M,N) \times M \to N$, $(f,x) \mapsto f(x)$ is continuous by Lemma B.10.

2.3 Lemma *Let L, O be manifolds and $h\colon L \to M$ and $f\colon N \to O$ be smooth. Then the pushforward and the pullback*

$$f_*\colon C^\infty(M,N) \to C^\infty(M,O), \quad g \mapsto f \circ g,$$
$$h^*\colon C^\infty(M,N) \to C^\infty(L,N), \quad g \mapsto g \circ h$$

are continuous.

Proof Since the compact open C^∞-topology is initial with respect to the family $(T^k)_{k\in\mathbb{N}_0}$, it suffices to prove that $T^k \circ f_*$ and $T^k \circ h^*$ are continuous for each $k \in \mathbb{N}_0$. However, the chain rule yields for each k commutative diagram

$$\begin{array}{ccc} C^\infty(M,N) & \xrightarrow{f_*} & C^\infty(M,O) \\ \downarrow{\scriptstyle T^k} & & \downarrow{\scriptstyle T^k} \\ C(T^kM, T^kN)_{\text{c.o.}} & \xrightarrow{(T^kf)_*} & C(M,O)_{\text{c.o.}}, \end{array} \qquad \begin{array}{ccc} C^\infty(M,N) & \xrightarrow{h^*} & C^\infty(L,N) \\ \downarrow{\scriptstyle T^k} & & \downarrow{\scriptstyle T^k} \\ C(T^kM, T^kN)_{\text{c.o.}} & \xrightarrow{(T^kh)^*} & C(T^kL, T^kN)_{\text{c.o.}} \end{array}$$

The pushforward and the pullback in the lower row are continuous by Lemma B.8. We conclude that f_* and h^* are continuous. □

2.4 Proposition *Let E be a locally convex space. Then the compact open C^∞-topology turns $C^\infty(M,E)$ with the pointwise operations into a locally convex space.*

Proof The compact open C^∞-topology is initial with respect to the map

$$\Phi\colon C^\infty(M,E) \to \prod_{k\in\mathbb{N}_0} \underbrace{C(T^kM, T^kE)_{\text{c.o.}}}_{\cong C(T^kM, E^{2^k})_{\text{c.o.}}}, \quad f \mapsto (T^kf)_{k\in\mathbb{N}_0}.$$

Now the spaces $C(T^kM, E^{2^k})_{\text{c.o.}}$ are locally convex spaces by Lemma B.7 since E^{2^k} is a locally convex space. The product of locally convex spaces is again a locally convex space. Thus every linear subspace of the product becomes a locally convex space. Now it is easy to see that Φ is linear with respect to pointwise addition and scalar multiplication. Thus the image of Φ is a linear subspace and Φ is an isomorphism of locally convex spaces identifying $C^\infty(M,E)$ as a locally convex subspace of $\prod_{k\in\mathbb{N}_0} C(T^kM, E^{2^k})$. □

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Interlude: Certain Open Sets in the Compact Open C^∞-topology

In this section, we recall the classic arguments (see e.g. Hirsch, 1994) that certain subsets of mappings are open in the compact open C^∞-topology. Observe that one needs here (and we shall require it in the whole section) for the source manifold to be compact. For non-compact source manifolds, the sets discussed here will, in general, not be open in the compact open C^∞-topology. We will need the results collected in this subsection in Chapter 3 when discussing the group of diffeomorphisms.

2.5 Lemma *Let M, N be manifolds, M compact and a finite open cover $U_1, \dots, U_n$ of M. Then the map*

$$\Psi\colon C^\infty(M,N) \to \prod_{i=1}^{n} C^\infty(U_i,N), \qquad f \mapsto (f|_{U_i})_{i\in I}$$

is a homeomorphism onto

$$\mathcal{I} := \{(f_i) \in \prod_{i=1}^{n} C^\infty(U_i,N) \mid f_i|_{U_i\cap U_j} \equiv f_j|_{U_i\cap U_j}, \text{ for all } i,j \in I\}.$$

Moreover, $\mathcal{I}$ is closed in $\prod_{i=1}^n C^\infty(U_i,N)$.

Proof To see that $\mathcal{I}$ is closed, we introduce, for every $i,j \in I := \{1,\dots,n\}$ and $x \in U_i \cap U_j$, the map $\mathrm{ev}_{x,i,j}\colon \prod_{i=1}^n C^\infty(U_i,N) \to N \times N$, $(\gamma_k)_{1\le k\le n} \mapsto (\gamma_i(x),\gamma_j(x))$. These are continuous since projections onto components in a product and the point evaluations are continuous (see Remark B.5). Now we denote by $\Delta N \subseteq N \times N$ the diagonal (i.e. all elements of the form (n,n)), which is closed in $N \times N$ due to N being Hausdorff. Then $\mathcal{I}$ is closed as the preimage:

$$\mathcal{I} = \bigcap_{x\in\bigsqcup_{i,j\in I} U_i\cap U_j} \mathrm{ev}_{x,i,j}^{-1}(\Delta N).$$

Note that we can write the restriction $f \mapsto f|_{U_i}$ as the pullback $f|_{U_i} = f \circ \iota_{U_i} = (\iota_{U_i})^*(f)$ with the inclusion of U_i into M. Hence the restriction map is continuous by Lemma 2.3, and as a consequence Ψ is continuous. Moreover, Ψ is clearly injective and we only have to prove that Ψ is an open mapping onto its image. To see this, we need to check that finite intersections of sets of the form

$$\lfloor K,U,k\rfloor := \{f \in C^\infty(M,N) \mid T^k f(K) \subseteq U\},$$
$$K \subseteq T^k M \text{ compact}, \quad U \subseteq\!\!\subseteq T^k N, \quad k \in \mathbb{N}$$

are mapped to open sets by Ψ (recall that the topology is initial with respect to the T^k). Now since M is compact, Lang (1999, II, §3 Proposition 3.2) implies

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

that there are open sets $V_i \subseteq \overline{V}_i \subseteq U_i, 1 \leq i \leq n$ and $M \subseteq \bigcup_{1\leq i\leq n} V_i$. Consider now $f \in \bigcap_{1\leq r\leq \ell} \lfloor K_r, U_r, k_r \rfloor$ and define $T^{k_r}\overline{V}_i := \pi_{k_r}^{-1}(\overline{V}_i)$, where $\pi_{k_r} : T^{k_r}U_i \to U_i$ is the bundle projection. Note that $T^{k_r}\overline{V}_i$ is closed in $T^{k_r}U_i$ for every $1 \leq i \leq n$. Then clearly

$$\gamma|_{U_i} \in \bigcap_{1\leq r\leq \ell} \lfloor K_r \cap T^{k_r}\overline{V}_i, O_r, k_r \rfloor \quad \text{for all } 1 \leq i \leq n. \tag{2.1}$$

Now let $(g_i)_{1\leq i\leq n} \in \mathcal{I}$ such that every g_i satisfies (2.1) for i. Since the $T^{k_r}V_i$ cover K_r, the unique map g defined by $g|_{U_i} = g_i$ satisfies $g \in \bigcap_{1\leq r\leq \ell} \lfloor K_r, U_r, k_r \rfloor$. We conclude

$$\prod_{i=1}^{n} \bigcap_{1\leq r\leq \ell} \lfloor K_r \cap T^{k_r}\overline{V}_i, O_r, k_r \rfloor \subseteq \Psi\left(\bigcap_{1\leq r\leq \ell} \lfloor K_r, U_r, k_r \rfloor \right)$$

and thus Ψ is open onto its image. □

2.6 Lemma *Let M be a compact manifold and N a finite-dimensional manifold. Then the sets*

$$\mathrm{Imm}(M,N) = \{f \in C^\infty(M,N) \mid f \textit{ is an immersion}\},$$
$$\mathrm{Sub}(M,N) = \{f \in C^\infty(M,N) \mid f \textit{ is a submersion}\}$$

are open in the compact open C^∞-topology.

Proof Since M is compact (hence finite dimensional), a map is an immersion (submersion) if and only if it is infinitesimally injective (surjective). We need to check that these properties define an open set in $C(TM,TN)$, whence they induce an open set in the compact open C^∞-topology.

Consider a map $f \in C^\infty(M,N)$ and pick a pair of charts (U_φ,φ) of M and (U_κ,κ) of N such that $f(U_\varphi) \subseteq U_\kappa$. In addition, we pick a compact set $K_\varphi \subseteq U_\varphi$ with non-empty interior. By compactness of M, we can choose a finite set of pairs of charts and compact sets such that the interior of the K_φ cover M. Apply Lemma 2.5 to obtain an embedding $C^\infty(M,N) \to \prod_{i=1}^n C^\infty(U_{\varphi_i},N)$. We will now construct for each K_{φ_i} an open neighbourhood in $C^\infty(U_{\varphi_i},N)$ which consists only of immersions (submersions) if f has this property. Pulling back the product of these neighbourhoods with the embedding then yields the desired neighbourhood of f in $C^\infty(M,N)$.

To this end, we consider the smooth map $g := \kappa \circ f \circ \varphi^{-1} \in C^\infty(V_\varphi,V_\kappa)$ where $V_\varphi \Subset \mathbb{R}^d$ and $V_\kappa \Subset \mathbb{R}^n$. Set $L := \varphi(K_\varphi)$ and observe that g is an immersion (submersion) if and only if f is an immersion (submersion). Recall that on open subsets of vector spaces we have $Tg = (g, dg) \in C(TV_\varphi, TV_\kappa) = C(V_\varphi \times \mathbb{R}^d, V_\kappa \times \mathbb{R}^n)$. Let $e_1,\dots,e_d$ be the standard basis of $\mathbb{R}^d$. For $x \in L$ we represent the Jacobian as $J_x(g) = [df(x;e_1),\dots,df(x,e_d)]$. If g is

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

an immersion (submersion) then the Jacobi matrix has for every x maximal rank, that is, if g is an immersion, the differential is injective if and only if $\mathrm{rk} J_x(g) = d \leq n$. In particular, the rank of the Jacobi matrix is constant, say, $\mathrm{rk} J_x(g) = N$ for all $x \in L$. We can thus pick for every $x \in L$ a subset $I_x \subseteq \{1,\dots,d\}$ of N elements such that $\{dg(x;e_j)\}_{j\in I_x}$ is linearly independent (note that the indices will, in general, depend on x!). If $\{df(x;e_j)\}_{j\in I_x}$ is linearly independent, then there exists $\varepsilon_x > 0$ such that every tuple $(x_1,\dots,x_N) \in \prod_{j\in I_x} B_\varepsilon(dg(x;e_j))$ is linearly independent (where $B_\varepsilon(z)$ is the ε-ball in $\mathbb{R}^d$; see Margalef-Roig and Domínguez, 1992, Lemma 1.6.7). By continuity of dg, there is for every $x \in L$ a compact neighbourhood N_x of x such that $(dg(y;e_{j_1}),\dots,dg(y;e_{j_N})) \in \prod_{j\in I_x} B_\varepsilon(dg(x;e_j)$ for all $y \in \overline{N}_x$. Thus

$$dg \in \Omega(g,x) := \bigcap_{j\in I_x} \lfloor N_x \times \{e_j\}, B_{\varepsilon_x}(df(x;e_j))\rfloor. \tag{2.2}$$

By construction, every $h \in C^\infty(V_\varphi,V_\kappa)$ with $dh \in \Omega(g,x)$ has a Jacobian of rank N at every point in N_x. In other words every such h is an immersion (submersion) on N_x if g is such a map. In Exercise 2.1.4 we shall now construct from $\Omega(g,x)$ an open neighbourhood of f in $C^\infty(U_\varphi,N)$ consisting only of maps which restrict to immersions (submersions) on K_φ if f is an immersion (submersions). We conclude that $\mathrm{Imm}(M,N)$ and $\mathrm{Sub}(M,N)$ are neighbourhoods of their points, hence open. □

2.7 Proposition *Let M be a compact manifold and N a finite-dimensional manifold. Then the set of embeddings*

$$\mathrm{Emb}(M,N) := \{f \in \mathrm{Imm}(M,N) \mid f \text{ is a topological embedding}\}$$

is open in the compact open C^∞-topology.

Proof Let $f \in \mathrm{Emb}(M,N)$ and fix a finite family of charts

(a) (U_i,φ_i) of M and (V_i,ψ_i) of N such that $f(U_i) \subseteq V_i$, and such that
(b) for every i there is a compact set $K_i \subseteq U_i$ and the interiors of the K_i cover M.

Recall that an embedding is, in particular, an injective immersion. Hence Lemma 2.6 allows us to choose an open neighbourhood $O_f \subseteq C^\infty(M,N)$ of f consisting only of immersions which satisfy also property (a). We shall now show that we can shrink O_f to obtain an open neighbourhood of f consisting only of immersions.

We have already seen in Lemma 1.50 that every immersion in O_f restricts locally to an embedding. However, since M,N are finite-dimensional manifolds, we can use the quantitative version of the inverse function theorem (see

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Glöckner, 2016, 1.1 and the references there) to obtain a uniform estimate on the size of these neighbourhoods: Shrinking O_f, we may assume that every $g \in O_f$ satisfies

(c) $g|_{U_i}$ is an embedding for every i

(an alternative proof of this fact using uniform estimates can be found in Hirsch, 1994, 2. Lemma 1.3). Now since f is an embedding, we see that for every i the compact sets $f(K_i)$ and $f(M \setminus U_i)$ are disjoint and we can thus find disjoint $A_i, B_i \subseteq N$ such that $f(K_i) \subseteq A_i$ and $f(M \setminus U_i) \subseteq B_i$. As there are only finitely many i, we can shrink O_f further such that every $g \in O_f$ satisfies

(d) $g(K_i) \subseteq A_i$ and $g(M \setminus U_i) \subseteq B_i$ for all i.

We shall now show that $g \in O_f$ is injective. Let $x, y \in M$ be distinct points and $x \in K_i$. If $y \in U_i$, then $g(x) \neq g(y)$ by (c). If $y \in M \setminus U_i$, then $g(x) \in A_i$ and $g(y) \in B_i$ by (d), so again $g(x) \neq g(y)$. We conclude that g is injective.

Summing up, O_f is an open set consisting entirely of injective immersions. However, since M is compact, every injective immersion is an embedding (see e.g. Margalef-Roig and Domínguez, 1992, Proposition 3.3.4). We conclude that O_f consists only of embeddings, whence $\mathrm{Emb}(M,N)$ is open. □

2.8 Corollary *If M is a compact manifold, the set of diffeomorphisms*

$$\mathrm{Diff}(M) := \{f \in C^\infty(M,M) \mid \text{there exists } g \in C^\infty(M,M) \text{ with } g \circ f = \mathrm{id}_M\}$$

is open in $C^\infty(M,M)$ with the compact open C^∞-topology.

Proof A diffeomorphism ϕ permutes the connected components of M and induces on every component a diffeomorphism onto another component. Since the components are compact, there is an open ϕ-neighbourhood in $C^\infty(M,M)$ whose elements map every component to the same component as ϕ. Thus we may assume that M is connected.

A diffeomorphism $\phi\colon M \to M$ is, in particular, a map which is an embedding and a submersion. Assume conversely that $\psi\colon M \to M$ is a mapping which is a submersion and an embedding. Since the image of a submersion is open (Exercise 1.7.5), the set $\phi(M)$ is open and closed in M, whence $\phi(M) = M$ by connectedness of M. Hence ϕ is a bijective map. Its inverse is smooth by Exercise 1.7.6 as ϕ is a submersion and $\mathrm{id}_M = \phi^{-1} \circ \phi$ (alternatively a bijective embedding) is a diffeomorphism by Lemma 1.61. □

2.9 Remark As was already mentioned for non-compact M, the subsets considered in this subsection will, in general, not form open subsets of $C^\infty(M,N)$ with respect to the compact open C^∞-topology. The reason for this is that the

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

compact open topology can only control a function's behaviour on compact sets. On a non-compact manifold, this topology is too weak to control the behaviour of a function on all of M. For this reason, one has to refine the topology if M is non-compact. The Whitney-type topologies are a common choice; see Hjelle and Schmeding (2017). However, many results presented in the next sections do not hold (at least not in the generality stated) for the Whitney-type topologies. A few examples of this behaviour for M non-compact are:

- the pullback h^* is, in general, discontinuous for the Whitney-topologies (whereas it is continuous in the compact open topology by Lemma 2.3);
- the exponential law, Theorem 2.12, is wrong.

While one can develop a general theory for function spaces on non-compact manifolds (see e.g. Michor, 1980), these examples show already that the resulting theory will require a much higher technical investment. We refrain from a discussion in the context of this book and refer the interested reader instead to the literature (Hjelle and Schmeding, 2017; Michor, 1980).

Exercises

2.1.1 Fill in the details for Remark 2.2.

(a) Show that the compact open C^∞-topology is the initial topology with respect to $(T^k)_{k\in\mathbb{N}}$.
Hint: A mapping into a product is continuous if and only if each component is continuous.

(b) If M, N are open subsets of locally convex spaces, show that the initial topologies with respect to the families $(T^k)_{k\in\mathbb{N}_0}$ and $(d^k)_{k\in\mathbb{N}_0}$ coincide.
Hint: Exercise 1.6.3 yields one inclusion of topologies. For the converse show inductively that $d^k \circ f$, for all $k \in \mathbb{N}_0$ continuous, implies $T^k \circ f$ continuous for all $f\colon Z \to C^\infty(M,N)$.

2.1.2 Let $\varphi\colon M \to N$ be a smooth map between smooth manifolds and E a locally convex space. Show that the pullback $\varphi^*\colon C^\infty(N,E) \to C^\infty(M,E)$, $f \mapsto f\circ\varphi$ is continuous linear. Deduce that if φ is a diffeomorphism, then φ^* is an isomorphism of locally convex spaces.

2.1.3 Let K, L be compact manifolds and M be a manifold. Show that the composition map $\mathrm{Comp}\colon C^\infty(K,M)\times C^\infty(L,K) \to C^\infty(L,M)$, $(f,g)\mapsto f\circ g$ is continuous.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

2.1.4 Fill in the missing details in the proof of Lemma 2.6. Show, in particular, that, thanks to the compactness of L, the $\Omega(g,x)$ yield an open neighbourhood of g in $C^\infty(V_\varphi, V_\kappa)$ consisting only of mappings whose Jacobian has maximal rank on the compact set L. Moreover, construct a neighbourhood of $f \in C^\infty(M,N)$ consisting entirely of immersions (submersions) if f is an immersion (submersion).

2.2 The Exponential Law and Its Consequences

In this section, we prove a version of the exponential law, Theorem 2.12, for smooth mappings. Before we begin, let us observe a crucial fact about the compact open C^∞-topology.

Assume that M is a *compact* manifold and E a locally convex space. Since the compact open C^∞-topology is finer than the compact open topology, we see that for every

$$O \subseteq E, \qquad C^\infty(M,O) := \{f \in C^\infty(M,E) \mid f(M) \subseteq O\}$$

is an open subset. Now $C^\infty(M,E)$ is a locally convex space by Proposition 2.4, whence $C^\infty(M,O)$ becomes a manifold and it makes sense to consider differentiable mappings with values in $C^\infty(M,O)$.

We now prepare the proof of the exponential law by providing several auxiliary results.

2.10 Lemma *Let E,F,H be locally convex spaces, $U \subseteq E$ and $V \subseteq F$. If $f: U \times V \to H$ is smooth, then so is $f^\vee: U \to C^\infty(V,H)$ $f^\vee(u) := f(u,\cdot)$. Its derivative is given by*

$$df^\vee(x;\cdot) = (d_1 f)^\vee(x). \tag{2.3}$$

Proof

Step 1: $f^\vee$ is continuous.

It suffices to prove that $d^k \circ f^\vee: U \to C(V \times F^k, H)_{\text{c.o.}}$ is continuous for every $k \in \mathbb{N}_0$ (see Remark 2.2 and Exercise 2.1.1). For $k = 0$ this was proved in Proposition B.13. We prove by induction that

$$d^k \circ f^\vee(x) = d^k(f^\vee(x)) = (d_2^k f)^\vee(x) \quad \text{for all } x \in U. \tag{2.4}$$

The induction start for $k = 0$ is trivial. For the induction step let $k > 0$ and we compute

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$\begin{aligned}
&d^k(f^\vee(x))(y;v_1,\dots,v_k)\\
&\quad=\lim_{t\to 0} t^{-1}\Big(d^{k-1}(f^\vee(x))(y+tv_k;v_1,\dots,v_{k-1})\\
&\qquad -d^{k-1}(f^\vee(x))(y;v_1,\dots,v_{k-1})\Big)\\
&\quad=\lim_{t\to 0} t^{-1}\Big(d^{k-1}(f(x,y+tv_k;(0,v_1),\dots,(0,v_{k-1}))\\
&\qquad -d^{k-1}(f(x,y;(0,v_1),\dots,(0,v_{k-1}))\Big)\\
&\quad=d^k f(x,y;(0,v_1),\dots,(0,v_k)).
\end{aligned}$$

Thus we have identified $d^k\circ f$ as $(d_2^k f)^\vee$ which is again continuous by Proposition B.13. We conclude that $f^\vee$ is continuous.

Step 2: $f^\vee$ is C^1 and the derivative satisfies (2.3).
Pick $x\in U$, $z\in E$ and $t\in\mathbb{R}$ small. We shall show that

$$\Delta(t,x,z):=t^{-1}(f^\vee(x+tz)-f^\vee(x))\xrightarrow{t\to 0}(d_1 f)^\vee(x,\cdot;z)$$

in $C^\infty(V,H)$. Recall that the compact open C^∞-topology is initial with respect to the family $(d^k\colon C^\infty(V,H)\to C(V\times F^k,H)_{\mathrm{c.o}})_{k\in\mathbb{N}_0}$. Thus $\Delta(t,x,z)$ converges for $t\to 0$ if and only if $d^k\circ\Delta(t)$ converges. Therefore, we pick $k\in\mathbb{N}_0$ and a neighbourhood $\lfloor K,U\rfloor$ of $d^k((d_1f)^\vee)$, i.e. $K\subseteq V\times F^k$ is compact such that $(d_1f)^\vee(x,\cdot,z)(K)\subseteq U\Subset H$. Since higher differentials are symmetric by Schwarz' theorem (Exercise 1.3.3), we have

$$\begin{aligned}
d^k(d_1f)^\vee(x;z)(y;v_1,\dots,v_k)&=d^{k+1}f(x,y;(z,0),(0,v_1),\dots,(0,v_k))\\
&=d^{k+1}f(x,y;(0,v_1),\dots,(0,v_k),(z,0)).
\end{aligned}$$

Now, we saw in Lemma 1.21 that the difference quotient extends continuously to $t=0$ by the differential. We apply this to d^kf: For each $\overline{y}:=(y_0,v_1,\dots,v_k)\in K\subseteq V\times F^k$, there exists $\overline{y}\in N_{\overline{y}}\Subset V\times F^k$ and $\varepsilon_{\overline{y}}>0$ such that

$$\begin{aligned}
&N_{\overline{y}}\times]-\varepsilon_{\overline{y}},\varepsilon_{\overline{y}}[\setminus\{0\}\to H,\\
&(\overline{w},t)\mapsto t^{-1}(d^kf(x+tz,w_0;(0,w_1),\dots,(0,w_k))\\
&\qquad\qquad -d^kf(x,w_0;(0,w_1),\dots,(0,w_k)),
\end{aligned}$$

where $\overline{w}=(w_0,\dots,w_k)$ the function takes values in U and extends continuously to some function $N_{\overline{y}}\times]-\varepsilon_{\overline{y}},\varepsilon_{\overline{y}}[\to U$. Exploiting compactness of K we cover it by finitely many of these neighbourhoods $N_{\overline{y}_1},\dots,N_{\overline{y}_\ell}$. Hence if $|t|<\min_{1\le i\le\ell}\varepsilon_{\overline{y}_i}$, we see that

$$t^{-1}(d^kf(x+tz,v;(0,v_1),\dots,(0,v_k))-d^kf(x,v;(0,v_1),\dots,(0,v_k))\in U.$$

In other words, $d^n\Delta(t,x,z)(\overline{v})\in U$ for all t small enough and $\overline{v}\in K$, hence (2.3) holds. Since $(d_1f)^\vee$ is C^0 by Step 1, we see that $f^\vee$ is C^1.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Step 3: f is C^k for each $k \geq 2$.
Note that $h\colon (U \times E) \times V \to H$, $((x,z),y) \mapsto d_1 f(x,y;z)$ is smooth. Now we argue by induction, where Step 2 is the induction start. Also Step 2 shows that $df^\vee = h^\vee$. The induction hypothesis shows that $h^\vee$ is C^{k-1}, so $f^\vee$ must be C^k and since k was arbitrary, $f^\vee$ is smooth. □

2.11 Proposition *Let E be a finite-dimensional space, F be locally convex and $U \subseteq E$. Then the evaluation map* $\operatorname{ev}\colon C^\infty(U,F) \times U \to F$ *is smooth.*

Proof We already know from Remark 2.2 that ev is continuous. Moreover, ev is linear in the first component and thus $d_1 \operatorname{ev}(f,x;g) = \operatorname{ev}(g,x)$ (this implies, in particular, that all partial derivatives with respect to the first component exist). Let us now compute $d_2 \operatorname{ev}$. For $f \in C^\infty(U,F)$ and $x \in U$, $w \in E$ and small t, we have

$$t^{-1} \operatorname{ev}(f, x+tw) - \operatorname{ev}(f,x) = t^{-1}(f(x+tw) - f(x)) \to df(x;w) \text{ as } t \to 0.$$

Hence $d_2 \operatorname{ev}(f,x;w)$ exists and is given by

$$d_2 \operatorname{ev}(f,x;w) = df(x;w) = \operatorname{ev}_1(df,(x,w)),$$

where $\operatorname{ev}_1\colon C^\infty(U \times E,F) \times (U \times E) \to F$, $(\gamma,z) \mapsto \gamma(z)$ is continuous. We conclude that

$$d \operatorname{ev}(f,x;g,v) = d_1 \operatorname{ev}(f,x;g) + d_2 \operatorname{ev}(f,x;v) = \operatorname{ev}(g,x) + \operatorname{ev}_1(df,(x,v)) \tag{2.5}$$

exists and is continuous (the mapping $C^\infty(U,F) \to C^\infty(U \times E,F)$, $f \mapsto df$ is clearly continuous and linear hence smooth). Thus ev (and also ev_1) is C^1. We see that inductively ev is C^k as its derivative is already C^{k-1}. □

We will now formulate and prove the exponential law, Theorem 2.12. To justify the name, denote the set of all functions from X to Y by Y^X. In this notation, the exponential law for arbitrary maps becomes $(Z^Y)^X \cong Z^{X \times Y}$, hence its name.

2.12 Theorem (Exponential law) *Let M be a compact manifold and $O \subseteq E, U \subseteq F$ be open subsets of locally convex spaces. Then*

(a) *If $f\colon U \times M \to O$ is smooth, so is $f^\vee\colon U \to C^\infty(M,O)$, $\quad f^\vee(x)(y) := f(x,y)$.*
(b) *The evaluation map* $\operatorname{ev}\colon C^\infty(M,O) \times M \to O$, $(\gamma,x) \mapsto \gamma(x)$ *is smooth and the map $C^\infty(U \times M,O) \to C^\infty(U,C^\infty(M,O))$, $f \mapsto f^\vee$ is a bijection.*

In particular, the mapping f is smooth if and only if $f^\vee$ is smooth.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Proof Since $C^\infty(M,O)$ is an open submanifold of $C^\infty(M,E)$, Lemma 1.39 shows that for the purpose of this proof we may assume without loss of generality that $O = E$.

(a) Pick a finite atlas $\{(\varphi_i, U_i)\}_{1\le i\le n}$ of the compact manifold M and obtain from Lemma 2.5 a topological embedding with closed image

$$\Psi\colon C^\infty(M,E) \to \prod_{i=1}^{n} C^\infty(U_i,E), \quad f \mapsto (f|_{U_i})_{1\le i\le n}.$$

Now since the image of Ψ is closed we apply Lemma 1.25: For any smooth $f\colon U\times M \to O$, the map $f^\vee\colon U \to C^\infty(M,O) \subseteq C^\infty(M,E)$ will be smooth if and only if $\Psi\circ f^\vee$ is smooth. In other words, $f^\vee$ is smooth if and only if for every $1\le i\le n$ the map $U \ni x \to f^\vee(x)|_{U_i} \in C^\infty(U_i,E)$ is smooth. Applying Exercise 2.1.2, this is equivalent to the smoothness of the mappings $U \ni x \to f^\vee(x)|_{U_i}\circ\varphi_i^{-1} \in C^\infty(\varphi_i(U_i),E)$ and these mappings are smooth by Lemma 2.10.

(b) The map $\mathrm{ev}\colon C^\infty(M,O)\times M \to O$ is smooth if locally around each point it is smooth. Hence we choose $(f,x)\in C^\infty(M,O)\times M$ and pick $\varphi\colon U\to V\subseteq \mathbb{R}^d$ a chart of M around x. We obtain another evaluation map $\mathrm{ev}_\varphi\colon C^\infty(V,O)\times V\to O$ such that

$$\mathrm{ev}(\eta,z) = \mathrm{ev}_\varphi((\varphi^{-1})^*(\eta),\varphi(z)), \quad (\eta,z)\in C^\infty(M,O)\times U.$$

Note that $(\varphi^{-1})^*\colon C^\infty(M,O)\to C^\infty(V,O)$ is the restriction of the continuous linear (hence smooth) map $(\varphi^{-1})^*\colon C^\infty(M,E)\to C^\infty(V,E)$ to the open set $C^\infty(M,O)$ (see Exercise 2.1.2). Hence $(\varphi^{-1})^*$ is smooth and ev will be smooth if ev_φ is smooth for each chart φ of M. However, the smoothness of ev_φ was established in Proposition 2.11.

We finally have to check that $C^\infty(U\times M,O)\to C^\infty(U,C^\infty(M,O))$, $f\mapsto f^\vee$ is bijective. Obviously it is injective; hence we need to check surjectivity. If $f\colon U\to C^\infty(M,O)$ is smooth, then the following map is smooth:

$$U\times M \to C^\infty(M,O)\times M, \quad (u,m)\mapsto (f(u),m).$$

Composing with ev, the map $f^\wedge\colon U\times M\to O$, $(u,m)\mapsto f(u)(m)$ is smooth and satisfies $(f^\wedge)^\vee = f$. This establishes surjectivity. □

2.13 Remark One can even show that the map from Theorem 2.12(b) is a homeomorphism. Moreover, one can obtain similar results for finite orders of differentiability. We skip the details here and refer instead to Alzaareer and Schmeding (2015) for more information.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Note that compactness of the manifold M is a crucial ingredient and the statement of the exponential law becomes false for general non-compact M; see Michor (1980).

Exercises

2.2.1 Let M be compact and E be a finite-dimensional vector space. Show that the evaluation map $\mathrm{ev}\colon C^\infty(M,E)\times M\to E$ is a submersion. (If you fancy a real challenge, prove this for a locally convex space E.)

2.2.2 Let $f\colon U\times M\to O$ and $p\colon O\to N$ be smooth maps. As always, we denote by $p_*\colon C^\infty(M,O)\to C^\infty(M,N)$, $h\mapsto p\circ h$, the push-forward and assume that the exponential law, Theorem 2.12, holds for the spaces $C^\infty(U\times M,O)$ and $C^\infty(U\times M,N)$. Prove that

$$p_*\circ(f^\vee)=(p\circ f)^\vee.$$

2.3 Manifolds of Mappings

In this section, we discuss spaces of smooth mappings between manifolds as infinite-dimensional manifolds. We shall not directly construct the manifold structure for general spaces (see Appendix C for a sketch).

General Assumption In this section K will be a compact smooth manifold, M,N will be smooth (possibly infinite-dimensional) manifolds.

2.14 Definition A smooth manifold structure on $C^\infty(K,M)$ is *canonical* if

- the underlying topology is the compact open C^∞-topology, and
- for each (possibly infinite-dimensional) C^∞-manifold N and for a map $f\colon N\to C^\infty(K,M)$, said map is C^∞ if and only if

$$f^\wedge\colon N\times K\to M,\quad (x,y)\mapsto f(x)(y)\quad \text{is } C^\infty.$$

2.15 Remark A canonical manifold structure enforces a suitable version of the exponential law, Theorem 2.12. This enables differentiability properties of mappings to be verified on the underlying manifolds.

A similar notion of canonical manifold exists also for spaces of finitely often differentiable mappings (Amiri et al., 2020; Glöckner and Schmeding, 2022). We hasten to remark that the usual constructions of manifolds of mappings yield canonical manifold structures (see Appendix C).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

2.16 Lemma *If $C^\infty(K,M)$ is endowed with a canonical manifold structure, then*

(a) *The evaluation map* $\mathrm{ev}\colon C^\infty(K,M)\times K\to M$, $\mathrm{ev}(\gamma,x):=\gamma(x)$ *is a C^∞-map.*
(b) *Canonical manifold structures are unique: writing $C^\infty(K,M)'$ for $C^\infty(K,M)$ with another canonical manifold structure, then*

$$\mathrm{id}\colon C^\infty(K,M)\to C^\infty(K,M)',\qquad \gamma\mapsto\gamma$$

is a C^∞-diffeomorphism.
(c) *Let $N\subseteq M$ be a submanifold such that the set $C^\infty(K,N)$ is a submanifold of $C^\infty(K,M)$. Then the submanifold structure on $C^\infty(K,N)$ is canonical.*
(d) *If M_1 and M_2 are smooth manifolds such that $C^\infty(K,M_1)$ and $C^\infty(K,M_2)$ have canonical manifold structures, then the manifold structure on the product manifold $C^\infty(K,M_1)\times C^\infty(K,M_2)\cong C^\infty(K,M_1\times M_2)$ is canonical.*

Proof (a) Since $\mathrm{id}\colon C^\infty(K,M)\to C^\infty(K,M)$ is C^∞ and $C^\infty(K,M)$ is endowed with a canonical manifold structure, it follows that $\mathrm{id}^\vee\colon C^\infty(K,M)\times K\to M$, $(\gamma,x)\mapsto\mathrm{id}(\gamma)(x)=\gamma(x)=\mathrm{ev}(\gamma,x)$ is C^∞.

(b) The map $f:=\mathrm{id}\colon C^\infty(K,M)\to C^\infty(K,M)'$ satisfies $f^\wedge=\mathrm{ev}$ and is thus C^∞, by (a). Since $C^\infty(K,M)'$ is endowed with a canonical manifold structure, it follows that f is C^∞. By the same reasoning, $f^{-1}=\mathrm{id}\colon C^\infty(K,M)'\to C^\infty(K,M)$ is C^∞.

(c) As $C^\infty(K,N)$ is a submanifold, the inclusion $\iota\colon C^\infty(K,N)\to C^\infty(K,M)$, $\gamma\mapsto\gamma$ is C^∞ (see Lemma 1.39). Likewise, the inclusion map $j\colon N\to M$ is C^∞. Let L be a manifold and $f\colon L\to C^\infty(K,N)$ be a map. If f is smooth, then $\iota\circ f$ is smooth, entailing that $(\iota\circ f)^\wedge\colon L\times K\to M$, $(x,y)\mapsto f(x)(y)$ is C^∞. As the image of this map is contained in N, which is a submanifold of M, we deduce that $f^\wedge=(\iota\circ f)^\wedge|^N$ is C^∞. If, conversely, $f^\wedge\colon L\times K\to N$ is C^∞, then also $(\iota\circ f)^\wedge=j\circ(f^\wedge)\colon L\times K\to M$ is C^∞. Hence $\iota\circ f\colon L\to C^\infty(K,M)$ is C^∞ (the manifold structure on the range being canonical). As $\iota\circ f$ is a C^∞-map with image in $C^\infty(K,N)$ which is a submanifold of $C^\infty(K,M)$, we deduce that f is C^∞.

(d) If L is a manifold and $f=(f_1,f_2)\colon L\to C^\infty(K,M_1)\times C^\infty(K,M_2)$ a map, then f is C^∞ if and only if f_1 and f_2 are C^∞. As the manifold structures are canonical, this holds if and only if $f_1^\wedge\colon L\times K\to M_1$ and $f_2^\wedge\colon L\times K\to M_2$ are C^∞, which holds if and only if $f^\wedge=(f_1^\wedge,f_2^\wedge)$ is C^∞. □

2.17 Proposition *Assume that $C^\infty(K,M)$ and $C^\infty(K,N)$ admit canonical manifold structures. If $\Omega\subseteq K\times M$ is an open subset and $f\colon\Omega\to N$ is a*

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

C^∞-map, then $\Omega' := \{\gamma \in C^\infty(K,M) \mid \{(k,\gamma(k)), k \in K\} \subseteq \Omega\}$ is an open subset of $C^\infty(K,M)$ and

$$f_\star : \Omega' \to C^\infty(K,N), \qquad \gamma \mapsto f \circ (\mathrm{id}_K, \gamma)$$

is a C^∞-map.

Proof In Exercise B.2.1 it was proved that Ω' is open in $C(K,M)_{\mathrm{c.o.}}$, and so it is open in the finer compact open C^∞-topology. By Lemma 2.16(a), the evaluation map $\mathrm{ev}: C^\infty(K,M) \times K \to M$ is C^∞, whence $C^\infty(K,M) \times K \to K \times M$, $(\gamma, x) \mapsto (x, \gamma(x))$ is C^∞. Since f is C^∞, the chain rule shows that

$$(f_\star)^\wedge : \Omega' \times K \to N, \qquad (\gamma, x) \mapsto f_\star(\gamma)(x) = f(x, \gamma(x)) = f(x, \mathrm{ev}(\gamma, x))$$

is C^∞. So $f_\star$ is C^∞, as the manifold structure on $C^\infty(K,N)$ is canonical. □

2.18 Corollary *Assume that $C^\infty(K,M)$ and $C^\infty(K,N)$ admit canonical manifold structures. If $f: K \times M \to N$ is a C^∞-map, then we obtain a smooth map*

$$f_\star : C^\infty(K,M) \to C^\infty(K,N), \qquad \gamma \mapsto f \circ (\mathrm{id}_K, \gamma).$$

Applying Corollary 2.18 with $f(x,y) := g(y)$, we get the following.

2.19 Corollary *Assume that $C^\infty(K,M)$ and $C^\infty(K,N)$ admit canonical manifold structures. If $g: M \to N$ is a C^∞-map, then the* pushforward *is smooth*

$$g_* : C^\infty(K,M) \to C^\infty(K,N), \qquad \gamma \mapsto g \circ \gamma.$$

To construct manifold structures on $C^\infty(K,M)$ one needs an additional structure on M. This so-called *local addition* replaces the vector space addition not present on M.

2.20 Definition Let M be a smooth manifold. A *local addition* is a smooth map

$$\Sigma : U \to M,$$

defined on an open neighbourhood $U \subseteq TM$ of the *zero-section* of the tangent bundle $\mathbf{0}_M := \{0_p \in T_pM \mid p \in M\}$ such that $\Sigma(0_p) = p$ for all $p \in M$,

$$U' := \{(\pi_M(v), \Sigma(v)) \mid v \in U\}$$

is open in $M \times M$ and $\theta := (\pi_{TM}, \Sigma): U \to U'$ is a diffeomorphism.

If $C^\infty(M,N)$ is canonical and we interpret a tangent vector as an equivalence class of smooth curves $[t \to c(t)]$, with $c:]-\varepsilon, \varepsilon[\to C^\infty(M,N)$, the derivative of c can be identified with the partial derivative of the adjoint map $c^\wedge :]-\varepsilon, \varepsilon[\times M \to N$. This shows that as a set we should have $TC^\infty(M,N) \cong C^\infty(M, TN)$. In the presence of a local addition, the set $C^\infty(M,TN)$ turns also

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

into a canonical manifold and the bijection becomes an isomorphism of vector bundles over the identity. Summing up, this identification yields the following result.

2.21 Proposition *If M admits a local addition*[1]*, then $C^\infty(K,M)$ admits a canonical manifold structure and the tangent bundle can be identified with $C^\infty(K,TM)$.*

We refer to Appendix C for more information about the proof.

2.22 Assume that M,N admit local additions and $f\colon M \to N$ is a C^∞-map. Then the identification $TC^\infty(K,M) \cong C^\infty(K,TM)$ induces a commuting diagram (Exercise 2.3.6):

$$\begin{array}{ccc} TC^\infty(K,M) & \xrightarrow{\cong} & C^\infty(K,TM) \\ \downarrow{\scriptstyle T(f_*)} & & \downarrow{\scriptstyle (Tf)_*} \\ TC^\infty(K,N) & \xrightarrow{\cong} & C^\infty(K,TN). \end{array} \tag{2.6}$$

We have seen in Corollary 2.19 and Exercise 2.3.1 that the pushforward and the pullback of smooth functions are smooth with respect to canonical manifolds of mappings. Viewing these mappings as partial mappings of the full composition map

$$\mathrm{Comp}\colon C^\infty(K,M)\times C^\infty(L,K) \to C^\infty(L,M), \quad (f,g)\mapsto f\circ g,$$

we see that the full composition is separately smooth in its variables. This immediately prompts the question as to whether the full composition map is smooth. In the general case (of a possibly non-compact source manifold) when one has no exponential law available this is complicated, but in our situation it reduces to an easy observation.

2.23 Proposition *Let K,L be compact manifolds and assume that the manifolds $C^\infty(K,M), C^\infty(L,M)$ are canonical ($C^\infty(L,K)$ is automatically a canonical manifold as L admits a local addition). Then the composition map*

$$\mathrm{Comp}\colon C^\infty(K,M)\times C^\infty(L,K) \to C^\infty(L,M), \qquad (f,g)\mapsto f\circ g$$

is smooth.

[1] One can show that every paracompact strong Riemannian manifold (see Chapter 4), and thus every finite-dimensional paracompact manifold, admits a local addition. Moreover, Lie groups (see Chapter 3) admit local additions, C.2.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Proof By the exponential law for canonical manifolds, Comp is smooth if and only if the adjoint map

$$\mathrm{Comp}^\wedge : C^\infty(K,M) \times C^\infty(L,K) \times L \to M, \quad (f,g,l) \mapsto f(g(l))$$

is smooth. However, this shows that $\mathrm{Comp}^\wedge(f,g,l) = \mathrm{ev}(f,\mathrm{ev}(g,l))$ and since the evaluation mappings are smooth for canonical manifolds, also the adjoint map and thus the composition are smooth. □

We have already seen that certain properties 'lift to the manifold of mappings'. For example, if $f\colon M \to N$ is smooth, the pushforward $f_*\colon C^\infty(K,M) \to C^\infty(K,N)$, $g \mapsto f \circ g$ is smooth. Another example of this is the following result whose proof is remarkably involved and technical (we omit the proof here and pose it as Exercise 2.3.7).

2.24 Lemma (Stacey–Roberts Lemma (Amiri and Schmeding, 2019, Lemma 2.4)) *Let $p\colon M \to N$ be a submersion between finite-dimensional manifolds. Endowing the function spaces with their canonical manifold structure, the pushforward $p_*\colon C^\infty(K,M) \to C^\infty(K,N)$ becomes a submersion.*

In the following chapters, we will study other structures from differential geometry which can be lifted from finite dimensions to spaces of differentiable functions. For Lie groups this leads to the so-called current groups (whose most prominent examples are the loop groups). In the context of Riemannian geometry, the lifting procedure gives rise to the L^2-metric and more generally to the Sobolev type Riemannian metrics on function spaces. Some examples of Sobolev type metrics will be discussed in §5.1 and Chapter 7.

Exercises

2.3.1 Let $h\colon L \to K$ be a smooth map. Assume that $C^\infty(K,M)$ and $C^\infty(L,M)$ are canonical manifolds.

(a) Show that the pullback $h^*\colon C^\infty(K,M) \to C^\infty(L,M)$, $f \mapsto f \circ h$ is smooth.

(b) Assume that K, L are compact and M admits a local addition. Then we identify $TC^\infty(K,M) \cong C^\infty(K,TM)$ (see C.12). Show that this identifies $T(h^*)$ with $h^*\colon C^\infty(K,TM) \to C^\infty(L,TM)$.

2.3.2 Let K be a compact manifold and $O \subseteq E$ in a locally convex space. Prove that $C^\infty(K,O) \subseteq C^\infty(K,E)$ (Proposition 2.4) is a canonical manifold.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

2.3.3 Let M be a finite-dimensional manifold and K be a compact manifold. Endow $C^\infty(K,M)$ with the canonical manifold structure from Appendix C.

(a) Show that for $x \in K$ the point evaluation $\mathrm{ev}_x \colon C^\infty(K,M) \to M$, $\gamma \mapsto \gamma(x)$ is a submersion.[2]

(b) Deduce that the set $S(x,y) := \{f \in C^\infty(K,M) \mid f(x) = y\}$ for some fixed $x \in K$, $y \in M$ is a split submanifold of $C^\infty(K,M)$.

(c) Is the set $\bigcap_{1\le i\le n} S(x_i,y_i)$ also a submanifold of $C^\infty(K,M)$ if we pick points $x_i \in K$, $y_i \in M$ for $1 \le i \le n$ and $n \in \mathbb{N}$?

2.3.4 Consider a compact manifold K and M a manifold with a local addition. We endow $C^\infty(K,M)$ with the canonical manifold structure induced by the local addition; see Appendix C.3. Compute the tangent map of the evaluation map

$$\mathrm{ev} \colon C^\infty(K,M) \times K \to M, \quad (\varphi, m) \mapsto \varphi(m).$$

Hint: Apply the rule on partial differentials, Exercise 1.6.3. To compute the derivative for the variable in $C^\infty(K,M)$, exploit the fact that $TC^\infty(K,M) \cong C^\infty(K,TM)$; see C.12. After choosing a smooth curve $c \colon]-\varepsilon,\varepsilon[\to C^\infty(K,M)$, apply the exponential law to carry out the computation.

2.3.5 Assume that K, L are compact and M admits a local addition. Compute a formula for the tangent map of the smooth map (see Proposition 2.23)

$$\mathrm{Comp} \colon C^\infty(L,M) \times C^\infty(K,L) \to C^\infty(K,M), \quad (g,f) \mapsto g \circ f.$$

2.3.6 Use the identification $TC^\infty(M,N) \ni [t \mapsto c] \mapsto (x \mapsto \frac{\partial}{\partial t} c^\wedge(t,x)) \in C^\infty(M,TN)$ to establish the commutativity of the diagram (2.6).

2.3.7 Let K be a compact manifold and $p \colon M \to N$ a submersion between finite-dimensional paracompact manifolds. Establish the Stacey–Roberts Lemma by showing that the pushforward $p_* \colon C^\infty(K,M) \to C^\infty(K,N)$ becomes a submersion.

Hint: This is an involved exercise in *finite-dimensional* geometry which should only be attempted if one is familiar with Riemannian exponential maps, parallel transport and horizontal distributions. The

[2] It is also possible to prove that the evaluation map $\mathrm{ev} \colon C^\infty(K,M) \times K \to M$, $(\gamma, x) \mapsto \gamma(x)$ is a submersion; see Schmeding and Wockel (2016, Corollary 2.9).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

idea is to construct a horizontal distribution $\mathcal{H}$ together with local additions η_M, η_N (constructed from suitable Riemannian exponential maps) such that the following diagram commutes:

$$\begin{array}{ccccc} TM = \mathcal{V} \oplus \mathcal{H} & \xleftarrow[\supseteq]{} & \Omega_M & \xrightarrow{\eta_M} & X \\ \downarrow{\scriptstyle 0 \oplus Tp|_{\mathcal{H}}} & & & & \downarrow{\scriptstyle p} \\ TN & \xleftarrow[\supseteq]{} & \Omega_N & \xrightarrow{\eta_N} & M. \end{array}$$

Using these local additions, the canonical charts of the manifold of mappings become submersion charts.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

3

Lifting Geometry to Mapping Spaces I: Lie Groups

In this chapter, one aim is to study spaces of mappings taking their values in a Lie group. It will turn out that these spaces carry again a natural Lie group structure. However, before we prove this, let us recall the definition and basic properties of (infinite-dimensional) Lie groups.

3.1 (Infinite-Dimensional) Lie Groups

Our presentation of Lie groups modelled on infinite-dimensional spaces mostly follows Neeb (2006). There are many accounts in the literature for finite-dimensional Lie theory (see, e.g. Hilgert and Neeb, 2012), but infinite-dimensional Lie theory (beyond Banach spaces) is by comparison relatively young and in its modern form goes back to the seminal work of Milnor (1982, 1984).

3.1 Definition A *(locally convex) Lie group* G is a manifold G modelled on a locally convex space endowed with a group structure such that the multiplication map $m_G : G \times G \to G$ and the inversion map $\iota : G \to G$ are smooth. A morphism of Lie groups is a smooth group homomorphism. In the following we shall drop the adjective 'locally convex' and simply say 'Lie group'.

Standard Notation Let us fix some standard notation for objects occurring frequently in conjunction with Lie groups. Let G be a Lie group, and we shall write

- $\mathbf{1}_G$ for the unit element (or shorter $\mathbf{1}$),
- m_G for multiplication, ι_G for inversion,
- for $g \in G$ we let $\lambda_g : G \to G$, $h \mapsto gh$, and $\rho_g : G \to G$, $h \mapsto hg$, the *left-(right-)translation*. (Observe that $\lambda_g(\rho_h(x)) = gxh = \rho_h(\lambda_g(x))$.)

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

3.2 Example A locally convex space E is a Lie group with respect to vector addition and the usual manifold structure.

3.3 Example The following examples are the classical finite-dimensional examples encountered in a first course on Lie theory. We include them here for readers who are not familiar with Lie groups.

(a) Let $M_n(\mathbb{R})$ be the set of all $n\times n$ matrices and $\mathrm{Gl}_n(\mathbb{R}):=\{A\in M_n(\mathbb{R})|\det A\neq 0\}$ be the set of $n\times n$ invertible matrices. Using the determinant, one sees that $\mathrm{Gl}_n(\mathbb{R})\subseteq M_n(\mathbb{R})\cong\mathbb{R}^{n^2}$ is an open subset, hence a manifold. Since multiplication of matrices is given by polynomials in the entries of matrices, the multiplication is smooth with respect to the manifold structure. For invertible matrices, Cramer's rule shows that inversion is also polynomial in the entries of the matrix, hence smooth. In conclusion, matrix multiplication and inversion turns $\mathrm{Gl}_n(\mathbb{R})$ into a Lie group.[1]

(b) The *orthogonal group* $\mathrm{O}_n(\mathbb{R}) := \{A\in\mathrm{Gl}_n(\mathbb{R}) \mid AA^\top = \mathrm{id}_{R^n}\}$ is a closed submanifold of $\mathrm{Gl}_n(\mathbb{R})$ and this structure turns it into a Lie group. Further, the *special orthogonal group* $\mathrm{SO}_n(\mathbb{R}) := \{A\in\mathrm{O}_n(\mathbb{R}) \mid \det(A)=1\}$ is an open subset of the orthogonal group and thus also a Lie group.

(c) The unit circle $\mathbb{S}^1\subseteq\mathbb{R}^2$ is a submanifold (as the unit sphere of the Hilbert space $\mathbb{R}^2$). Identifying $\mathbb{R}^2\cong\mathbb{C}$, complex multiplication induces a Lie group structure on $\mathbb{S}^1$ which is explicitly given by the formulae

$$(x,y)\cdot(a,b) := (xa-yb, xb+ay), \quad (x,y)^{-1} := (x,-y).$$

3.4 Example (Unit groups of continuous inverse algebras) To generalise the matrix group example to infinite dimensions, we recall the notion of a *continuous inverse algebra (CIA)*: Let A be a locally convex space with a continuous bilinear map $\beta\colon A\times A\to A$ (we write shorter $xy := \beta(x,y)$) such that the associativity law $(xy)z = x(yz)$ holds. Furthermore, we assume that there exists an element $\mathbf{1}\in A$ such that $\mathbf{1}\,x = x = x\,\mathbf{1}$ for all $x\in A$, and define the set of all invertible elements

$$A^\times := \{x\in A \mid \text{ there exists } x^{-1}\in A \text{ such that } xx^{-1} = \mathbf{1} = x^{-1}x\}.$$

Then $A^\times$ is a group, under the multiplication, called the *unit group* of A. If $A^\times$ is open in A and inversion $\iota\colon A^\times\to A^\times$, $x\mapsto x^{-1}$ is continuous, we call A a *continuous inverse algebra (CIA)*. The unit group of a CIA is a Lie group. See Exercise 3.1.2.

[1] Alternatively, this example can be seen as a special case of the unit group of a continuous inverse algebra; see Example 3.4.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

CIAs generalise Banach algebras, for example, the algebra of continuous linear operators $L(E,E)$ of a Banach space E with the operator norm $\|\cdot\|_{\mathrm{op}}$ is a CIA.

Before we continue with the general infinite-dimensional Lie theory, we will discuss now one of the most important example classes of such Lie groups: the diffeomorphism groups. These groups will return as a running example in the later sections to illustrate the concepts of Lie algebra and regularity.

3.5 Example Let M be a compact manifold. Then M possesses a local addition, whence $C^\infty(M,M)$ is a canonical manifold and by Proposition 2.23 the composition map $\mathrm{Comp}\colon C^\infty(M,M)\times C^\infty(M,M)\to C^\infty(M,M)$ is smooth. Recall from Corollary 2.8 that the *set of diffeomorphisms* $\mathrm{Diff}(M)$ is an open subset of $C^\infty(M,M)$ which forms a group under composition of smooth maps. Hence $\mathrm{Diff}(M)$ is an open submanifold of $C^\infty(M,M)$ and the group product is smooth with respect to this structure.

We will now prove that inversion $\iota\colon \mathrm{Diff}(M)\to\mathrm{Diff}(M)$ is smooth. Applying the exponential law for a canonical manifold, ι is smooth if and only if the mapping $\iota^\wedge\colon \mathrm{Diff}(M)\times M\to M$, $(\varphi,m)\mapsto\varphi^{-1}(m)$ is smooth. Consider the implicit equation

$$\mathrm{ev}(\phi,\iota^\wedge(\phi,m))=\phi(\iota^\wedge(\phi,m))=m,\quad \text{for all } m\in M, \tag{3.1}$$

which takes values in a finite-dimensional manifold, but has an infinite-dimensional parameter ($\phi\in\mathrm{Diff}(M)$). However, $\mathrm{ev}\colon \mathrm{Diff}(M)\times M\to M$ is smooth, and we can compute its partial differential as $T_{(\phi,x)}\,\mathrm{ev}(0,z)=T\phi(z)$ (cf. Exercise 2.3.4). Since ϕ is a diffeomorphism, we see that the partial derivative of ev is indeed invertible for every $\phi\in\mathrm{Diff}(M)$. Now smoothness of $\iota^\wedge$ follows from a suitable implicit function theorem. Observe that due to the infinite-dimensional parameter, the usual implicit function theorem Lang (1999, I. §5 Theorem 5.9) is not applicable to (3.1)! However, we invoke the generalised implicit function theorem (Glöckner, 2006b, Theorem 2.3) which can deal with parameters in locally convex spaces (as long as the target of the implicit equation is a Banach manifold). The usual application of the implicit function theorem to (3.1) shows then that $\iota^\wedge$ and thus ι is smooth. We conclude that $\mathrm{Diff}(M)$ is a Lie group.

3.6 Remark To establish smoothness of the inversion $\mathrm{Diff}(M)$, we needed the exponential law and a generalised implicit function theorem. Use of this machinery can be avoided: In Michor (1980, Theorem 11.11) differentiability of the inversion is directly verified (which is technical and requires the (non-trivial) verification of continuity first). Our approach is inspired by the proof

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

in the convenient setting (Kriegl and Michor, 1997, Theorem 43.1). There the problem can be reduced to a finite-dimensional equation which circumvents the need for a generalised implicit function theorem.

The Lie group $\mathrm{Diff}(M)$ already comes with a canonical action on the manifold M which we describe after recalling the notion of a Lie group action.

3.7 Definition (Lie group action) Let M be a manifold and G be a Lie group. Then a smooth map

$$\alpha\colon G \times M \to M, \quad (g,m) \mapsto \alpha(g,m) =: g \cdot m$$

is called a *(left) Lie group action* if it satisfies the following:

$$\alpha(\mathbf{1}_G, m) = m, \qquad \alpha(g_1, \alpha(g_2,m)) = \alpha(g_1 g_2, m), \quad \text{for all } g_1, g_2 \in G,\ m \in M.$$

A right action is a smooth map $\beta\colon M \times G \to M$, $(m,g) \mapsto \beta(g,m) =: m \cdot g$ such that

$$\beta(m, \mathbf{1}_G) = m, \qquad \beta(\beta(m,g_1),g_2) = \beta(m, g_1 g_2), \quad \text{for all } g_1, g_2 \in G,\ m \in M.$$

If β is a right action, then $\alpha := \beta \circ (\mathrm{id}_M, \iota_G)$ is a left action. Similarly, we can obtain right actions from left actions and there is no essential difference between the notions.

3.8 Example Let M be a compact manifold. Then, going through the construction in Example 3.5, it is immediately clear that the evaluation map

$$\alpha\colon \mathrm{Diff}(M) \times M \to M, \quad (\varphi, m) \mapsto \varphi(m)$$

induces a (left) Lie group action, called the *canonical action of the diffeomorphism group*. Furthermore, there is also the right action of $\mathrm{Diff}(M)$ on smooth functions

$$\beta\colon C^\infty(M,N) \times \mathrm{Diff}(M) \to C^\infty(M,N), \quad (f,\varphi) \mapsto \varphi^*(f) = f \circ \varphi.$$

If $C^\infty(M,N)$ is a canonical manifold of mappings, then Proposition 2.23 shows that β is a (right) Lie group action. The right action β is connected to several geometric structures such as the symplectic structure of the loop space $C^\infty(\mathbb{S}^1, N)$; see Wurzbacher (1995). We will encounter it again in the context of shape analysis in Chapter 5.

For a left Lie group action α the canonical map

$$\alpha^\vee\colon G \to \mathrm{Diff}(M), \quad g \mapsto \alpha(g,\cdot)$$

makes sense and yields a group morphism. If M is finite dimensional, the exponential law shows that smoothness of the group action is equivalent to smoothness of $\alpha^\vee$. This breaks down for an infinite-dimensional manifold M as there is no smooth structure on $\mathrm{Diff}(M)$. Similarly, if a Lie group G acts

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

by linear mappings on a vector space E (this is called a representation of G) the literature considers the smoothness of $\alpha^\vee$ as a mapping to $\mathrm{Aut}(E)$ (the group of linear automorphisms of E). If E is normable, the group $\mathrm{Aut}(E)$ inherits a canonical Lie group structure from the operator norm topology such that smoothness of α is equivalent to smoothness of $\alpha^\vee$. Again this equivalence breaks down for locally convex spaces which are not normable, as $\mathrm{Aut}(E)$ does, in general, not carry a Lie group structure. Note that this is not a serious problem, as one can still check smoothness with respect to the product $G \times M$ (and in infinite-dimensional representation theory of Lie groups even weaker concepts of smoothness of representations are more appropriate for the theory; see e.g. Neeb (2010) and also Neeb (2005, I.3.4)). However, we shall not discuss representation theory and the finer points of these problems in this book.

3.9 Definition Let G be a Lie group. We call a submanifold $H \subseteq G$ a *Lie subgroup* if it is a subgroup of G.[2] If H is, in addition, closed in G, we call H a closed Lie subgroup.

3.10 Example The diffeomorphism group $\mathrm{Diff}(M)$ of a compact manifold M contains many important subgroups of diffeomorphisms which preserve geometric structures. If ω is a differential form on M, we say *a diffeomorphism* $\phi \in \mathrm{Diff}(M)$ *preserves the differential form*, if $\phi^*\omega = \omega$ (where the pullback is as in Definition E.7). As the pullback commutes with function composition, we can consider the subgroup $\mathrm{Diff}_\omega(M)$ of diffeomorphisms preserving a given differential form ω. The most important examples are the following subgroups:

(a) if $\omega = \mu$ is a volume form on M, we obtain the group $\mathrm{Diff}_\mu(M)$ of *volume-preserving diffeomorphisms*;
(b) for a symplectic form ω, this yields the *group of symplectomorphism*;
(c) $\mathrm{Diff}_\theta(M) = \{\phi \in \mathrm{Diff}(M) \mid \phi^*\theta = f\theta \text{ for some } f \in C^\infty(M, \mathbb{R}\}$, the *group of contactomorphisms* for a contact form θ.

In the three cases mentioned above, one can show that the subgroups are also submanifolds of $\mathrm{Diff}(M)$ and thus Lie subgroups of $\mathrm{Diff}(M)$. We refer to Smolentsev (2007, Section 3) for detailed proofs. See, however, E.18 for a sketch of the construction for volume-preserving diffeomorphisms.

In finite-dimensional Lie theory a useful result states that every closed subgroup of a finite-dimensional Lie group is a Lie subgroup (Hilgert and Neeb, 2012, Theorem 9.3.2). This is no longer true in infinite dimensions, as the next example shows (see also Neeb, 2005, Remark V.2.4 (c)).

[2] Note that thanks to Lemma 1.39 this structure turns H into a Lie group.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

3.11 Example (Wockel, 2014) Consider the space $(\ell^2, +)$ of all real sequences which are square summable (Meise and Vogt, 1997, Example 12.11). This is a Hilbert space with respect to the inner product

$$\langle (x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} \rangle := \sum_{n\in\mathbb{N}} x_n y_n \qquad \left(\text{and norm } \|(x_n)_{n\in\mathbb{N}}\| = \sqrt{\sum_{n\in\mathbb{N}} x_n^2} \right).$$

We consider ℓ^2 as an abelian Lie group and define the subgroup

$$H := \left\{ (x_n)_{n\in\mathbb{N}} \in \ell^2 \mid x_n \in \frac{1}{n}\mathbb{Z},\ n \in \mathbb{N} \right\}.$$

As the projections $\pi_j \colon \ell^2 \to \mathbb{R}$, $(x_n)_{n\in\mathbb{N}} \mapsto x_j$, $j \in \mathbb{N}$, are continuous linear, $H = \bigcap_{n\in\mathbb{N}} \pi_n^{-1}(\frac{1}{n}\mathbb{Z})$ is a closed subgroup. However, we shall see in Exercise 3.1.5 that H is not a submanifold (it is not a manifold with respect to the subspace topology).

For Lie groups, the tangent bundle is again a Lie group and moreover, the tangent bundle is trivial (i.e. it splits as a product of a vector space and the base manifold).

3.12 Lemma *Let G be a Lie group. Identify $T(G\times G) \cong TG \times TG$.*

(a) *The tangent map of the multiplication*

$$\begin{aligned} Tm_G \colon T(G\times G) \cong TG\times TG &\to TG, \\ T_gG \times T_hG \ni (v_g, w_h) &\mapsto Tm_G(v_g, w_h) = T_g\rho_h(v_g) + T_h\lambda_g(w_h) \end{aligned} \tag{3.2}$$

induces a Lie group structure on TG with identity element $0_{\mathbf{1}} \in T_{\mathbf{1}}G$ and inversion

$$T\iota_G \colon TG \to TG, \qquad T_gG \ni v \mapsto -T\rho_{g^{-1}}T\lambda_{g^{-1}}(v) = -T\lambda_{g^{-1}}T\rho_{g^{-1}}(v). \tag{3.3}$$

The projection $\pi_G \colon TG \to G$ becomes a morphism of Lie groups with kernel $(T_{\mathbf{1}}G, +)$ and the zero-section $\mathbf{0} \colon G \to TG$, $g \mapsto 0_g$ is a morphism of Lie groups with $\pi_G \circ \mathbf{0} = \mathrm{id}_G$.

(b) *The map*

$$\Phi \colon G \times T_{\mathbf{1}}G \to TG, \qquad (g, v) \mapsto g \cdot v := Tm_G(0_g, v)$$

is a diffeomorphism.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Proof (a) Since m_G and ι_G are smooth, the same holds for their tangent maps. The group axioms for TG follow from the ones for G by virtue of the chain rule (and have the claimed unit element). Linearity of the tangent map implies (3.2); we leave this formula and (3.3) as Exercise 3.1.3. From the definition of the zero section and the projection, the morphism properties follow. As a result of (3.2), $T_{(\mathbf{1},\mathbf{1})}m_G(v_\mathbf{1}, w_\mathbf{1}) = v_\mathbf{1} + w_\mathbf{1}$. Hence the multiplication on the normal subgroup $\ker\pi_G = T_\mathbf{1}G$ is the addition.

(b) Since $\Phi = Tm_G(0_g, v) = Tm_G(\mathbf{0}(g), v)$ and the zero-section $\mathbf{0}$ is smooth, smoothness of the multiplication shows that Φ is smooth. Now a computation shows that $\Phi^{-1}(v) = (\pi_G(v), T_{\pi_G(v)}\lambda_{\pi_G(v)^{-1}}(v))$, whence Φ is bijective and its inverse is smooth (as inversion and multiplication in G are smooth and the projection π_G is smooth). □

3.13 Remark Note that Lemma 3.12 shows that $T_\mathbf{1}G$ is a normal Lie subgroup of TG, and TG is as a Lie group a semidirect product $T_\mathbf{1}G \rtimes G$.[3] Moreover, instead of left multiplication one can use right multiplication to identify the tangent bundle (the two different choices are related by the adjoint action; see Example 3.28).

The tangent space at the identity of a Lie group plays a special role. In the next section this tangent space will be endowed with an additional structure, the Lie bracket.

Exercises

3.1.1 Verify that $\mathbb{S}^1$ is a Lie group with the structure described in Example 3.3(c).

3.1.2 In this exercise we verify that the unit group $A^\times$ of a CIA (A, β) forms a Lie group. Note that the multiplication is smooth by Exercise 1.3.2. Hence it suffices to prove smoothness of inversion.

(a) Use the identity $b^{-1} - a^{-1} = b^{-1}(a - b)a^{-1}$ to deduce that the differential quotient $d\iota(x; y)$ exists and satisfies $d\iota(x; y) = -x^{-1}yx^{-1}$.

(b) Use the formula from (a) to prove that ι is C^1 and inductively is C^k for all $k \in \mathbb{N}_0$.

[3] A Lie group G with normal Lie subgroup N and Lie subgroup H such that $N \cap H = \{\mathbf{1}_G\}$ is a semidirect product $N \rtimes H$ if there exists a split exact sequence of Lie group homomorphisms

$$0 \to N \hookrightarrow G \xrightarrow{p} H \to \mathbf{1}$$

where splitting means that $p|_H = \mathrm{id}_H$. Equivalently, $G \cong N \times H$ (as manifolds) and the group product is given by $(n, h) \cdot (\tilde{n}, \tilde{h}) = (n\tilde{h}\tilde{n}\tilde{h}^{-1}, h\tilde{h})$. See Hilgert and Neeb (2012, 2.2.2) for further alternative characterisations.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

3.1.3 Fill in the missing details in the proof of Lemma 3.12.

(a) Prove (3.2) and verify that the tangent maps induce a group structure on TG.

(b) Establish (3.3), that is, $T_a\iota(v) = -T\lambda_{a^{-1}}T\rho_{a^{-1}}(v) = -T\rho_{a^{-1}}T\lambda_{a^{-1}}(v)$.
Hint: Let $\gamma\colon]-\varepsilon,\varepsilon[\to G$ be smooth with $\gamma(0) = a$. Differentiate the relation $\mathbf{1} = \gamma(t)(\gamma(t))^{-1}$.

(c) Show that one can obtain a diffeomorphism $TG \cong T_{\mathbf{1}}G\times G$ using right multiplication instead of left multiplication in part (b).

3.1.4 Let (A,β) be a continuous inverse algebra (CIA) and let $C^\infty(K,A)$ be endowed with the compact open C^∞-topology.

(a) Show that then $C^\infty(K,A)$ with the pointwise product is a CIA.

(b) Show that $C^\infty(K,A^\times) = C^\infty(K,A)^\times$ and thus the group $C^\infty(K,A^\times)$ with the pointwise product is a Lie group.

3.1.5 We supply the details for Example 3.11: Let $H = \{(x_n)_{n\in\mathbb{N}} \in \ell^2 \mid x_n \in \frac{1}{n}\mathbb{Z}, n \in \mathbb{N}\}$.

(a) Show that every 0-neighbourhood in the subspace topology of H contains at least one non-zero element.
Hint: It suffices to consider norm balls.

(b) Show that there is no 0-neighbourhood in H which contains a continuous path connecting 0 with a non-zero element. Deduce that H is not locally homeomorphic to an open subset of a locally convex space and thus is not a (sub)manifold.

3.1.6 Let G be a Lie group, and show that the map $L\colon G\times TG \to TG$, $(g,v_h)\mapsto T\lambda_g(v_h)$ is a left Lie group action. Dually right multiplication yields a right action $R\colon TG\times G\to TG$. Work out a formula relating $L(g,v_h)$ to $R(v_h,g)$.

3.2 The Lie Algebra of a Lie Group

We now associate a Lie algebra to a Lie group. This construction allows one to reformulate many problems in Lie theory in terms of linear algebra.

3.14 Definition A *Lie algebra* is a vector space $\mathfrak{g}$ together with a *Lie bracket*, that is, a bilinear map $[\cdot,\cdot]\colon \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ such that

(a) $[x,[y,z]]+[y,[z,x]]+[z,[x,y]] = 0$, for all $x,y,z\in\mathfrak{g}$ (Jacobi identity);
(b) $[x,x] = 0$, for all $x\in\mathfrak{g}$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

If $\mathfrak{g}$ is a locally convex space and the Lie bracket is continuous, $\mathfrak{g}$ is a *locally convex Lie algebra*. A (continuous) linear map $h\colon \mathfrak{g} \to \mathfrak{h}$ between (locally convex) Lie algebras is a morphism of (locally convex) Lie algebras if $h([x,y]) = [h(x), h(y)]$, for all $x, y \in \mathfrak{g}$.

3.15 Remark It will be essential for us that the Lie bracket and Lie algebra morphisms are continuous (see e.g. the proof of Proposition E.14 for an example where continuity is needed). From now on we will mostly work with locally convex Lie algebras; hence we drop the phrase 'locally convex' and write only 'Lie algebra'.

3.16 Example If A is a continuous inverse algebra (or more generally a locally convex algebra), then the commutator $[x,y] := xy - yx$ turns A into a Lie algebra. Hence the Lie bracket measures commutativity of the algebra product.

3.17 Example Every locally convex space E is a Lie algebra, called an *abelian Lie algebra*, with the *trivial bracket* $[x,y] := 0$.

3.18 (The Lie algebra of vector fields) Let M be a manifold, and let

$$\mathcal{V}(M) := \{X \in C^\infty(M, TM) \mid \pi_M \circ X = \mathrm{id}_M\}$$

be the locally convex space of all vector fields (see Appendix D). If $f \in C^\infty(M,E)$ is smooth with values in some locally convex space E and $X \in \mathcal{V}(M)$, then we obtain a smooth function

$$X.f := df \circ X\colon M \to E \quad (\text{recall } df = \mathrm{pr}_2 \circ Tf).$$

For $X, Y \in \mathcal{V}(M)$, there exists a unique vector field $[X,Y] \in \mathcal{V}(M)$ determined by the property that on each $U \subseteq M$ we have

$$[X,Y].f = X.(Y.f) - Y.(X.f) \quad \text{for all } f \in C^\infty(U,E). \tag{3.4}$$

Thus $\mathcal{V}(M)$ becomes a Lie algebra (the local case is verified in Appendix D and we discuss the general case in Exercise 3.2.3). If M is finite dimensional, Corollary D.13 shows that $(\mathcal{V}(M), [\cdot,\cdot])$ is a locally convex Lie algebra.

3.19 Let G be a Lie group. A vector field $X \in \mathcal{V}(G)$ is called *(left) invariant* if X is λ_g-related to itself for all $g \in G$ (i.e. $X \circ \lambda_g = T\lambda_g \circ X$; see Appendix D). We write $\mathcal{V}^\ell(G)$ for the set of left-invariant vector fields. Note that relatedness of vector fields is inherited by the Lie bracket thanks to Exercise 3.2.3, and so $\mathcal{V}^\ell(G)$ is a Lie subalgebra of $\mathcal{V}(G)$.

3.20 Proposition *Let G be a Lie group; then the map*

$$\Theta\colon T_{\mathbf{1}}G \to \mathcal{V}^\ell(G), \quad v \mapsto (g \mapsto T_{\mathbf{1}}\lambda_g(v))$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

is an isomorphism of locally convex spaces with inverse $\Theta^{-1}(X) = X(\mathbf{1})$. *Thus* $\mathbf{L}(G) := T_{\mathbf{1}}G$ *can be endowed with the Lie bracket*

$$[v, w] := \Theta^{-1}([\Theta(v), \Theta(w)]) = [\Theta(v), \Theta(w)](\mathbf{1}),$$

turning it into a Lie algebra. We call $(\mathbf{L}(G), [\cdot, \cdot])$ *the* Lie algebra associated to G.

Proof As $\Theta(v)(hg) = T_{\mathbf{1}}\lambda_{hg}(v) = T_g\lambda_h T_{\mathbf{1}}\lambda_g(v) = T_g\lambda_h\Theta(v)(g)$, the map Θ makes sense and its image consists of left-invariant vector fields (see Exercise 3.2.4). Linearity of Θ follows directly from the linearity of the tangent map. For $X \in \mathcal{V}^\ell(G)$ we have $X(g) = X \circ \lambda_g(\mathbf{1}) = T_{\mathbf{1}}\lambda_g X(\mathbf{1}) = \Theta(X(\mathbf{1}))(g)$, Θ is surjective. As the translations λ_g are diffeomorphisms, it is clear that only $0 \in T_{\mathbf{1}}G$ gets mapped to the zero-vector field. We conclude that Θ is a vector space isomorphism (its inverse is obviously evaluation in $\mathbf{1}$). Note that $\mathcal{V}^\ell(G)$ carries the subspace topology induced by $\mathcal{V}(G)$ from Appendix D.4. This immediately shows that Θ^{-1} is continuous since point evaluations are continuous in this topology. The continuity of Θ is left as Exercise 3.2.5. That $[\cdot, \cdot]$ is a Lie bracket on $T_{\mathbf{1}}G$ follows directly by trivial computations since $\mathcal{V}^\ell(G)$ is a Lie algebra. □

If the Lie group G is finite dimensional, the above discussion shows that $(T_{\mathbf{1}}G, [\cdot, \cdot])$ is a locally convex Lie algebra. Here only the continuity of the Lie bracket is unclear in the general case. We shall now prove that the Lie bracket on $T_{\mathbf{1}}G$ is always continuous; hence the Lie algebra $\mathbf{L}(G)$ associated to a Lie group is always a locally convex Lie algebra. To this end, we need a local model of the multiplication.

3.21 Let G be a Lie group. Since $T_{\mathbf{1}}G$ is isomorphic to the model space of G, we can pick a chart $\varphi\colon G \supseteq U_\varphi \to V_\varphi \subseteq T_{\mathbf{1}}G$ such that $\mathbf{1} \in U_\varphi$ and $\varphi(\mathbf{1}) = 0$. Moreover, we may assume that $T_{\mathbf{1}}\varphi = \mathrm{id}_{T_{\mathbf{1}}G}$. Due to the continuity of the multiplication of G there is an open $\mathbf{1}$-neighbourhood W with $W \cdot W \subseteq U_\varphi$ (here $W \cdot W$ denotes the set of all products of two elements in W). Hence we can define a local multiplication

$$*\colon \varphi(W) \times \varphi(W) \to V_\varphi, \quad (x, y) \mapsto x * y := \varphi(\varphi^{-1}(x)\varphi^{-1}(y)).$$

By construction, the local multiplication is smooth and $*(0, x) = x = *(x, 0)$. Hence the construction gives rise to a so-called local Lie group (see Neeb, 2005, Remark III.1.14). As with the Lie group G we can compute a Lie bracket using a local version of a left-invariant vector field. To distinguish the local operations from the Lie group operations, let us introduce a new symbol for

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

left translation: $\ell_x : \varphi(W) \to T_{\mathbf{1}}G$, $y \mapsto x * y$, $x \in \varphi(W)$. For any $v \in T_{\mathbf{1}}G$ we can thus define a left-invariant vector field with respect to the local product

$$\Lambda^v : \varphi(W) \to T_{\mathbf{1}}G, \quad x \mapsto d\ell_x(0; v) = \left.\frac{d}{dt}\right|_{t=0} x * tv.$$

We will see in Exercise 3.2.6 that a left-invariant vector field X with $X(\mathbf{1}) = v$ is φ-related to Λ^v. Together with the properties of the chart this yields the identity

$$\begin{aligned}[v,w] &= [\Theta(v),\Theta(w)](\mathbf{1}) = [\Lambda^v,\Lambda^w](0) = d\Lambda^w(0;\Lambda^v(0)) - d\Lambda^v(0;\Lambda^w(0)) \\ &= \left(\left.\frac{d^2}{dtds}\right|_{t,s=0} sv * tw - \left.\frac{d^2}{dtds}\right|_{t,s=0} tw * sv \right).\end{aligned}$$

This formula shows immediately that the Lie bracket $[\cdot,\cdot]$ is continuous on $T_{\mathbf{1}}G$.

3.22 Corollary *The Lie algebra* $(\mathbf{L}(G), [\cdot,\cdot])$ *associated to a Lie group is a locally convex Lie algebra.*

3.23 Remark (Left–right confusion) The reader may wonder now why one uses left-invariant vector fields to compute the Lie algebra. Instead one could as well use *right-invariant vector fields*, that is, $X(g) = T\rho_g X(\mathbf{1})$. This would also lead to a Lie algebra structure on $T_{\mathbf{1}}G$; however, the induced Lie bracket would have the opposite sign (see Exercise 3.2.7). Indeed there is no reason to prefer left-invariant vector fields over right-invariant ones: The choice of the former is historically motivated and customary.

3.24 Remark There are several alternative ways to introduce the Lie bracket on $T_{\mathbf{1}}G$. For example, one can use the adjoint action of the Lie group (which is briefly discussed in Example 3.28). In addition, the construction of the Lie bracket using the local multiplication in 3.21 can be interpreted (see Neeb, 2005, Lemma III.1.6) as a computation of the antisymmetric part of the second-order Taylor polynomial of the local multiplication at $(0,0)$. Moreover, the Lie bracket measures commutativity of the group multiplication (see Exercise 3.2.8).

3.25 Example (The Lie algebra of $\mathrm{Diff}(M)$) Let M be a compact manifold. In Example 3.5 we have seen that by identifying $\mathrm{Diff}(M)$ as an open subset of $C^\infty(M,M)$, it becomes a Lie group under composition of maps. Further, Proposition 2.21 shows that $T_{\mathrm{id}}\,\mathrm{Diff}(M) = T_{\mathrm{id}}C^\infty(M,M) \cong \mathcal{V}(M)$. We shall now show that the Lie algebra is given as $(\mathcal{V}(M), -[\cdot,\cdot])$ with the *negative* of the usual bracket of vector fields (see 3.18). Extend $X \in \mathcal{V}(M) = T_{\mathrm{id}}\ \mathrm{Diff}\ (M)$ to the right-invariant vector field $R_X \in \mathcal{V}^\rho(\mathrm{Diff}(M)), R_X(\varphi) := X \circ \varphi$ and

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

consider the product vector field $R_X \times \mathbf{0}_M \in \mathcal{V}(\mathrm{Diff}(M) \times M) = \mathcal{V}(\mathrm{Diff}(M)) \times \mathcal{V}(M)$. We now exploit the canonical action $\alpha\colon \mathrm{Diff}(M) \times M \to M$, $(\varphi, m) \mapsto \varphi(m)$, Example 3.8, and note that as a restriction of ev, a manifold version of (2.5) (see Exercise 2.3.4) yields the tangent map of α:

$$\begin{aligned} T\alpha(R_X \times \mathbf{0}_M)(\varphi, m) &= T_{(\varphi,m)}\alpha(X \circ \varphi, \mathbf{0}_M(m)) = X \circ \varphi(m) + T\varphi(0_m) \\ &= X \circ \varphi(m) = X(\alpha(\varphi, m)). \end{aligned}$$

Hence the product vector field $R_X \times \mathbf{0}_M$ is α-related to X. Thus for $X, Y \in \mathcal{V}(M)$ the bracket $[X, Y]$ is α-related to $[R_X \times \mathbf{0}_M, R_Y \times \mathbf{0}_M] = [R_X, R_Y] \times \mathbf{0}_M$, whence $[R_X, R_Y]$ is the negative of the usual bracket, see Remark 3.23.

If M admits a volume form μ, we have seen that the volume-preserving diffeomorphisms $\mathrm{Diff}_\mu(M)$ form a Lie subgroup of $\mathrm{Diff}(M)$. Thanks to E.18, the Lie algebra $\mathbf{L}(\mathrm{Diff}_\mu(M))$ of this subgroup can be identified as the Lie subalgebra of divergence-free vector fields $\mathcal{V}_\mu(M) = \{X \in \mathcal{V}(M) \mid \operatorname{div} X = 0 (\Leftrightarrow \mathcal{L}_X \mu = 0)\}$.

3.26 Example Consider the locally convex space E as the abelian Lie group $(E, +)$. An easy computation shows that $\mathbf{L}(E) = E$ with the trivial Lie bracket. Thus the Lie algebra of this abelian Lie group is an abelian Lie algebra (see Exercise 3.2.8).

We can associate to every Lie group morphism a morphism of (locally convex) Lie algebras as the following lemma shows.

3.27 Lemma *If $f\colon G \to H$ is a Lie group morphism (i.e. a smooth group homomorphism) then the map $\mathbf{L}(f) := T_{\mathbf{1}} f\colon \mathbf{L}(G) \to \mathbf{L}(H)$ is a morphism of Lie algebras, that is, $\mathbf{L}(f)([v, w]) = [\mathbf{L}(f)(v), \mathbf{L}(f)(w)]$, for all $v, w \in \mathbf{L}(G)$.*

Proof Let $v \in T_{\mathbf{1}} G$ and $\tilde{v} := T_{\mathbf{1}} f(v) \in T_{\mathbf{1}} H$. Since f is a group morphism we have

$$\Theta(\tilde{v})(f(g)) = T_{\mathbf{1}} \lambda_{f(g)} T_{\mathbf{1}} f(v) = T_g f(T_{\mathbf{1}} \lambda_g(v)) = T_g f \Theta(v).$$

Hence for every $v \in T_{\mathbf{1}} G$ the left-invariant vector field $\Theta(v)$ is f-related to $\Theta(\tilde{v})$. As f-relatedness is inherited by the Lie bracket (Exercise 3.2.4) we see that

$$T_{\mathbf{1}} f[\Theta(v), \Theta(w)] = [\Theta(\tilde{v}), \Theta(\tilde{w}] \circ f$$

evaluating in $\mathbf{1}$, we obtain the claimed formula since $f(\mathbf{1}_G) = \mathbf{1}_H$. □

3.28 Example Let G be a Lie group and $g \in G$. Then the conjugation with g is the Lie group morphism $c_g\colon G \to G$, $h \mapsto ghg^{-1}$. Hence for every g we obtain a Lie algebra morphism $\mathrm{Ad}_g := \mathbf{L}(c_g)\colon \mathbf{L}(G) \to \mathbf{L}(G)$, called the *adjoint map* of g. This gives rise to a smooth mapping

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$\mathrm{Ad}\colon G \times \mathbf{L}(G) \to \mathbf{L}(G), \quad (g,x) \mapsto \mathrm{Ad}_g(x),$$

called the *adjoint action* of G.

The upshot of this short repetition is that every Lie group comes with an associated Lie algebra, and every Lie group morphism gives rise to a Lie algebra morphism. In finite dimensions, the interplay between these objects leads to the classical Lie theorems.

The Lie Theorems for Finite-Dimensional Lie Groups and Lie Algebras

Lie 2 *If G,H are Lie groups and G is simply connected, then for every Lie algebra homomorphism $f\colon \mathbf{L}(G) \to \mathbf{L}(H)$ there exists a Lie group morphism with $f = \mathbf{L}(\varphi)$.*

Lie 3 *For every Lie algebra $\mathfrak{g}$ there exists a connected Lie group G with $\mathbf{L}(G) = \mathfrak{g}$.*

We omitted the first Lie theorem as it is a purely local statement which does not admit a global formulation on the Lie group. It is well known that the third Lie theorem fails in infinite dimensions. For example, one can show the following.

3.29 (A Lie algebra without an associated Lie group; Omori, 1981) Let M be a connected *non-compact* finite-dimensional manifold. Then there exists no Lie group G such that $\mathbf{L}(G) = \mathcal{V}(M)$.

However, under certain assumptions one can salvage at least the second Lie theorem. The main issue here is that in infinite-dimensional spaces, even simple differential equations might not have solutions (see Appendix A.6 for explicit counterexamples). Thus in our general setting, a major task is to establish the existence of solutions to differential equations relevant to Lie theory. These *equations of Lie type* will be discussed next.

Exercises

3.2.1 Let $(\mathfrak{g},[\cdot,\cdot])$ be a Lie algebra. Show that $[x,x] = 0$ for all $x \in \mathfrak{g}$ is equivalent to skew-symmetricity of the Lie bracket, that is, $[x,y] = -[y,x]$, for all $x,y \in \mathfrak{g}$.

3.2.2 Show that the commutator bracket (Example 3.16) is a Lie bracket.

3.2.3 With the help of the material in Appendix D show that the bracket of vector fields from 3.18 turns $\mathcal{V}(M)$ into a Lie algebra.

(a) Let $X,Y \in \mathcal{V}(M)$ and $\mathcal{A}$ be an atlas of M. The brackets of the local representatives yield $([X_\phi,Y_\phi])_{\phi\in\mathcal{A}}$. Show that this induces a vector field $[X,Y]$ with the properties from 3.18.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(b) Prove that $[X, X] = 0$ for all $X \in \mathcal{V}(M)$ and the Jacobi identity holds.

(c) Show that the bracket is continuous with respect to the topology from D.4.
Hint: All assertions can be localised in charts where Lemma D.11 holds.

3.2.4 Prove a global version of Lemma D.9, that is, if $(X_i, Y_i) \in \mathcal{V}(M) \times \mathcal{V}(N)$, $i = 1, 2$, are pairs of f-related vector fields, then $[X_1, X_2]$ and $[Y_1, Y_2]$ are f-related.

3.2.5 We check several details in the proof of Proposition 3.20. Show that:

(a) $X_v \colon G \to TG$, $g \mapsto T_{\mathbf{1}}\lambda_g(v)$ is a smooth left-invariant vector field for $v \in T_{\mathbf{1}}G$;

(b) $\Theta \colon T_{\mathbf{1}}G \to \mathcal{V}^\ell(M) \subseteq \mathcal{V}(M)$, $v \mapsto X_v$ is continuous.
Hint: Combine D.4 with Lemma 2.10.

3.2.6 Let G be a Lie group and φ be a chart with the properties from 3.21 with respect to which we define a local multiplication $*$. Let $X \in \mathcal{V}^\ell(G)$ be left invariant with $X(\mathbf{1}) = v$. Prove that $X_\varphi|_{\varphi(W)} = \Lambda^v$ and deduce that the (principal part of) left-invariant vector fields are thus locally related to the left-invariant vector fields with respect to the local multiplication.

3.2.7 Let $v \in \mathbf{L}(G)$ and denote by L_v, R_v the left- (/right-)invariant vector field constructed from v. Show that L_v and $-R_v$ are ι-related and deduce $[L_v, L_w] = -[R_v, R_w]$.
Hint: Use Exercise 3.1.3 to show that $T_{\mathbf{1}}\iota(v) = -v$.

3.2.8 Let G be an abelian Lie group. Show that the Lie bracket of $\mathbf{L}(G)$ is trivial, that is, the Lie bracket vanishes and the Lie algebra is abelian.
Hint: Since G is abelian all left-invariant vector fields are also right-invariant.

3.2.9 Let A be a CIA. Show that the Lie algebra of $A^\times$ is A with Lie bracket $[a, b] = ab - ba$.
Hint: Note that since $A^\times \subseteq A$, the equation for left invariance is $X_v(g) = gv (= \beta(g, v))$ in A.

3.2.10 Let $\mathfrak{g}, \mathfrak{h}$ be two Lie algebras. Show that there is a canonical way to turn the product $\mathfrak{g} \times \mathfrak{h}$ into a Lie algebra. Explain how this was exploited in Example 3.25.

3.2.11 Let G be a Lie group with associated Lie algebra $(\mathbf{L}(G).[\cdot, \cdot])$. In this exercise, we consider the adjoint action from Example 3.28. Show that:

(a) $\mathrm{Ad} \colon G \times \mathbf{L}(G) \to \mathbf{L}(G)$ is a Lie group action;

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(b) for $\varphi\colon G \to H$ a Lie group morphism, $\mathrm{Ad}_{\varphi(g)}(\mathbf{L}(\varphi)(x)) = \mathbf{L}(\varphi)(\mathrm{Ad}_g(x))$;
(c) for $x, y \in \mathbf{L}(G)$ one has $\mathrm{ad}_x(y) \coloneqq T_{\mathbf{1}_G,y}\,\mathrm{Ad}(x, 0_y) = [x, y]$.

3.3 Regular Lie Groups and the Exponential Map

In this section, we discuss differential equations needed for advanced tools in Lie theory.

3.30 Definition Let G be a Lie group with Lie algebra $\mathbf{L}(G)$. We say G is *semiregular* if for each smooth curve $\eta \in C^\infty([0,1], \mathbf{L}(G))$ the initial value problem

$$\begin{cases} \dot\gamma(t) &= \gamma(t).\eta(t) = T_{\mathbf{1}}\lambda_{\gamma(t)}(\eta(t)) \\ \gamma(0) &= \mathbf{1} \end{cases} \tag{3.5}$$

has a (unique) solution $\mathrm{Evol}(\eta) \coloneqq \gamma\colon [0,1] \to G$. We also say that (3.5) is a *Lie type equation*. The group G is *regular (in the sense of Milnor)* if G is semiregular and the following *evolution map* is smooth: (cf. C.15)

$$\mathrm{evol}\colon C^\infty([0,1], \mathbf{L}(G)) \to G, \quad \eta \mapsto \mathrm{Evol}(\eta)(1)$$

3.31 For a smooth curve $c\colon [a,b] \to G$, we can define the *left logarithmic derivative* $\delta^\ell(c)\colon [a,b] \to \mathbf{L}(G)$, $t \mapsto T\lambda_{c(t)}(\dot c(t))$. Note that the logarithmic derivative inverts the evolution, that is, $\delta^\ell(\mathrm{Evol}(\eta)) = \eta$. There are many identities relating δ^ℓ, Evol and their counterparts defined via right multiplication (see Kriegl and Michor (1997, Section 38) for an account, also Exercise E.2.3). The left- (right-)logarithmic derivative is closely connected to the left- (right-) Maurer–Cartan form on G; see Example E.10.

3.32 Remark (a) Solutions to (3.5) are automatically unique by Kriegl and Michor (1997, 38.3 Lemma).
(b) For regular Lie groups, Lie's second theorem holds. Since the proof can most conveniently be formulated within the framework of differential forms, we defer it to Appendix E.2.
(c) Instead of using the left multiplication in (3.5) one can also use right multiplication to define regularity. Similar to the definition of the Lie algebra, using inversion of the group shows that the two notions of regularity are the same.

3.33 Remark Thanks to the usual solution theory of ordinary differential equations, every Banach–Lie group (i.e. Lie group modelled on a Banach space) and thus every finite-dimensional Lie group is regular; see Neeb (2006).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

3.34 Example Consider the locally convex space E as a Lie group $(E,+)$. Note that its Lie algebra is again E with the zero-bracket (Exercise 3.3.3). For a smooth curve $\eta\colon [0,1] \to \mathbf{L}(E) = E$ we interpret the Lie type equation in $TE = E \times E$ and obtain $(\gamma(t),\gamma'(t)) = (\gamma(t),\eta(t))$. Hence a solution γ of (3.5) satisfies $\gamma' = \eta$. Therefore, if $(E,+)$ is regular, then E is Mackey complete; Definition 1.12. Conversely, if E is Mackey complete, then $\operatorname{evol}_E(\eta) := \int_0^1 \eta(s)\mathrm{d}s$ defines a continuous linear (hence smooth) map $\operatorname{evol}_E\colon C^\infty([0,1],E) \to E$, whence $(E,+)$ is regular if and only if it is Mackey complete. Note that one can show that any Lie group which is regular is necessarily modelled on a Mackey complete space (Neeb, 2006, Remark II.5.3(b)).

3.35 Example Let $(A,\cdot)$ be a continuous inverse algebra (CIA) which is Mackey complete. If the topology of A is generated by a family of seminorms which are submultiplicative, that is, $q(xy) \leq q(x)q(y)$, for all $x,y \in A$, then $(A^\times,\cdot)$ is regular (Glöckner and Neeb, 2012). In this case, the solutions to the evolution equation are given by the *Volterra series*

$$\gamma(t) = \mathbf{1} + \sum_{n\in\mathbb{N}} \int_0^t \int_0^{t_{n-1}} \cdots + \int_0^{t_2} \eta(t_1)\cdots\eta(t_n)\mathrm{d}t_1\cdots\mathrm{d}t_n. \tag{3.6}$$

This has interesting applications in physics, control theory and rough path theory, as for certain CIAs the above series models signatures of irregular paths; see Chapter 8.

3.36 Example ($\operatorname{Diff}(M)$ is regular) In Example 3.5 we saw that for a compact manifold M the group $\operatorname{Diff}(M)$ is a Lie group with Lie algebra $\mathcal{V}(M)$; Example 3.25. If $c\colon [0,1] \to \mathcal{V}(M) \subseteq C^\infty(M,TM)$ is smooth, we use the exponential law to interpret $X := c^\wedge$ as a smooth time-dependent vector field on M. Following Remark 3.32(c) we can solve Lie type equations with respect to the right multiplication to establish regularity. As right multiplication ρ_ϕ in $\operatorname{Diff}(M)$ is the pullback with ϕ, we deduce from Exercise 2.3.1 $T_{\mathbf{1}}\rho_\phi(c(t)) = c(t)\circ\phi$. Hence the exponential law allows us to rewrite the Lie type equation (3.5) with respect to right multiplication on $\operatorname{Diff}(M)$ as a differential equation for the vector field X_t (subscript denoting time dependence):

$$\gamma'(t) = X_t(\gamma(t)), \qquad \gamma(0) = \mathrm{id}_M. \tag{3.7}$$

In other words, γ solves the Lie type equation (3.5) (with respect to right multiplication) if and only if it satisfies (3.7), whence γ is the flow Fl^X of the time-dependent vector field X on $[0,1]\times M$ (see D.5). Since $[0,1]\times M$ is compact, the usual (finite-dimensional!) theory of ordinary differential equations (Lang, 1999, IV. §2) shows that the flow of such a vector field always exists. We deduce that $\operatorname{Diff}(M)$ is semiregular and can consider the evolution operator $\operatorname{evol}\colon C^\infty([0,1],\mathcal{V}(M)) \to \operatorname{Diff}(M)$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

To see that the evolution is smooth, recall that $\mathrm{Diff}(M) \subseteq C^\infty(M,M)$. Now we exploit that $C^\infty(M,M)$ is a canonical manifold of mappings to deduce that evol is smooth if

$$\mathrm{evol}^\wedge \colon C^\infty([0,1],\mathcal{V}(M)) \times M \to M, \quad (X,m) \mapsto \mathrm{Fl}_1^X(m)$$

is smooth. Here Fl_1^X denotes the time 1-flow of the time-dependent vector field X. We view (3.7) now as an ordinary differential equation (ODE) whose right-hand side $F(t,X,m) = \mathrm{ev}(X,t)(m) = X(t,m)$ depends smoothly on the parameter X. The theory of parameter-dependent ordinary differential equations shows that Fl^X depends smoothly on X; see, for example, Alzaareer and Schmeding (2015, Proposition 5.13), which permits the (infinite-dimensional!) space $C^\infty([0,1],\mathcal{V}(M))$ as a parameter space. Hence $\mathrm{evol}^\wedge$ is smooth and $\mathrm{Diff}(M)$ is regular.

As a consequence of the regularity of $\mathrm{Diff}(M)$, we can apply Lie's second theorem, Proposition E.14, to Lie algebra morphisms into the Lie algebra of vector fields on a compact manifold. Milnor (1982) used these observations to prove a restricted version of the Lie–Palais theorem; see Palais (1957). In its general form, the Lie–Palais theorem asserts that every finite-dimensional Lie algebra of vector fields $\mathfrak{g}$ of vector fields on a finite-dimensional smooth manifold M, which is generated by complete vector fields,[4] consists of complete vector fields and can be integrated to a global action of a Lie group G on M. We discuss a proof for a severely limited version of this result as Exercise 3.3.7.

Not all Lie groups are regular, as one can construct some pathological examples modelled on *incomplete* spaces. However, regularity has been established for almost all naturally occurring classes of Lie groups. Nevertheless, the following conjecture is still open.

Conjecture (Milnor, 1983) *Every Lie group modelled on a Mackey-complete space is regular.*

For regular Lie groups, one can solve the important class of Lie type differential equations. This allows us to discuss the Lie group exponential function (which can be viewed as an abstraction of the matrix exponential function to Lie groups).

3.37 Lemma *Let G be a regular Lie group and $X \in \mathcal{V}^\ell(G)$. Then there exists a unique curve $\gamma_X \colon \mathbb{R} \to G$ with $\dot\gamma_X(t) = X(\gamma_X(t))$, for all t and $\gamma_X(0) = \mathbf{1}$. This implies that every left-invariant vector field is complete.*

Proof Consider the smooth (constant) curve $\eta_X \colon \mathbb{R} \to T_{\mathbf{1}}G$, $t \mapsto X(\mathbf{1})$ (in the following we will frequently consider the restriction of η_X to $[0,1]$ without

[4] A vector field $X \in \mathcal{V}(M)$ is complete if its integral curves $\dot\varphi_x(t) = X(\varphi_x(t))$, $\varphi_x(0) = x$ exist for all $t \in \mathbb{R}$ and $x \in M$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

further notice). By regularity, we obtain a unique solution $\gamma\colon [0,1] \to G$ such that $\gamma(0) = \mathbf{1}$ and $\dot\gamma(t) = T\lambda_{\gamma(t)}(\eta_X(t)) = X(\gamma(t))$, $t \in [0,1]$. Thus this curve is the flow of the invariant vector field. Since the (constant) curve η_X makes sense for all $t \in \mathbb{R}$, we can consider the equation (3.5) for every $t \in \mathbb{R}$. We now extend the flow γ to all of $\mathbb{R}$.

Step 1: Smooth extension to $[-1,1]$.
Using right-regularity, we can construct a curve $\gamma_R\colon [0,1] \to G$ which satisfies $\gamma(0) = \mathbf{1}$ and $\dot\gamma(t) = T\rho_{\gamma(t)}(X(1))$. Let us show that the formula

$$\gamma_X(t) := \begin{cases} \gamma_R^{-1}(-t), & t \in [-1,0], \\ \gamma(t), & t \in [0,1] \end{cases}$$

yields a smooth extension of γ whose derivative is $X(\gamma_X(t))$ at every t. We compute with the formula for the derivative of the inverse (3.3), the derivative of γ_X on $[-1,0]$:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\gamma_R^{-1}(-t) &= -T\lambda_{\gamma_R^{-1}(-t)} T\rho_{\gamma_R^{-1}(-t)}\dot\gamma_R(-t) \\ &= T\lambda_{\gamma_R^{-1}(-t)} T\rho_{\gamma_R^{-1}(-t)} T\rho_{\gamma_R(-t)}(v) = X(\gamma_R^{-1}(-t)). \end{aligned}$$

Hence γ_X is a smooth integral curve of the left-invariant field X.

Step 2: γ_X *extends to all of* $\mathbb{R}$.
Pick $0 < t_0 < 1$ and define $\gamma_{t_0}\colon [-1+t_0, 1+t_0] \to G$, $t \mapsto \gamma(t_0)\gamma_X(t-t_0)$. Note that $\gamma_{t_0}(t_0) = \gamma_X(t_0)$ and furthermore,

$$\begin{aligned} \dot\gamma_{t_0}(t_0+t) &= T\lambda_{\gamma(t_0)}(\dot\gamma_X(t)) = T\lambda_{\gamma(t_0)}(X(\gamma_X(t))) = X(\gamma(t_0)\gamma_X(t)) \\ &= X(\gamma_{t_0}(t+t_0)), \end{aligned}$$

where we have used the left invariance of X. Uniqueness of the solution implies now that on their common domain of definition γ_X and γ_{t_0} coincide. Thus we can extend γ_X to $[-1, 1+t_0]$. Repeating the argument, the domain of γ_X is not bounded from above. Choosing $-1 < t_0 < 0$ a similar argument shows that the domain cannot be bounded from below. Thus γ_X can be continued for all of $\mathbb{R}$. □

3.38 Definition Let G be a regular Lie group. Then we define the *Lie group exponential*

$$\exp_G\colon \mathbf{L}(G) \to G, \quad v \mapsto \gamma_v(1),$$

where γ_v is the unique integral curve starting from $\mathbf{1}$ of the vector field $L_v(g) := T\lambda_g(v)$.

3.39 Remark Note that $C\colon \mathbf{L}(G) \to C^\infty([0,1], \mathbf{L}(G))$, $v \mapsto (t \mapsto v)$ is continuous linear, hence smooth. Thus $\exp_G = \mathrm{evol} \circ C$ is smooth for every regular Lie group G.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

3.40 Lemma *Let G be a regular Lie group. Define for $\eta \in C^\infty([0,1], \mathbf{L}(G))$ and $s \in [0,1]$ the curve $\eta_s(t) := \eta(st)$. Then*

$$\operatorname{Evol}(s\eta_s)(t) = \operatorname{Evol}(\eta)(st) \text{ for all } t \in [0,1] \quad \text{and} \quad \operatorname{Evol}(\eta)(s) = \operatorname{evol}(s\eta_s).$$

In particular, this implies $T_0 \exp = \mathrm{id}_{\mathbf{L}(G)}$.

Proof The curve $t \mapsto \operatorname{Evol}(\eta)(st)$ takes 0 to the identity in G, and its derivative is $\frac{\mathrm{d}}{\mathrm{d}t} \operatorname{Evol}(\eta)(st) = s \operatorname{Evol}(\eta)(st) \cdot \eta(st)$. Here we abuse the multiplication to denote the (derivative) of left translation. Thus it solves the Lie type equation for the curve $s\eta_s$. This proves the first identity, while we obtain the second for $t = 1$. Now let $\eta(t) := v$ be constant for $v \in \mathbf{L}(G)$. Then $\eta_s(t) = \eta(t)$ and we see that $\exp(sv) = \operatorname{evol}(s\eta_s) = \operatorname{Evol}(\eta)(s)$. Derivating at $s = 0$ yields

$$T_0 \exp(v) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \exp(sv) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \operatorname{Evol}(\eta)(s) = \underbrace{\operatorname{Evol}(\eta)(0)}_{=\mathbf{1}_G} \cdot \eta(0) = v.$$

□

Unfortunately, the observation that $T_0 \exp = \mathrm{id}_{\mathbf{L}(G)}$ is not as useful as in the Banach setting, where the inverse function theorem implies that the Lie group exponential is a local diffeomorphism onto a neighbourhood of the unit of the group (then the Lie group exponential yields a canonical chart called exponential coordinates). In general, the Lie group exponential need be *neither* locally injective *nor* locally surjective. Most prominently, this happens for the diffeomorphism group $\operatorname{Diff}(M)$ from Example 3.5. We discuss the special case for $M = \mathbb{S}^1$ in Example 3.42. Lie groups for which the exponential function is well behaved thus deserve a special name.

3.41 Definition A regular Lie group G is called *locally exponential* if the exponential function $\exp\colon \mathbf{L}(G) \to G$ restricts to a local diffeomorphism between a neighbourhood of $0 \in \mathbf{L}(G)$ and $\mathbf{1}_G \in G$.

Unfortunately, diffeomorphism groups are not locally exponential, as the following classical example shows.

3.42 Example Consider the unit circle $\mathbb{S}^1$. We will show that the image of the Lie group exponential of the diffeomorphism group $\operatorname{Diff}(\mathbb{S}^1)$ contains no identity neighbourhood. Recall from Example 3.36 that the Lie group exponential is the map

$$\exp\colon \mathbf{L}(\operatorname{Diff}(\mathbb{S}^1)) = \mathcal{V}(\mathbb{S}^1) \to \operatorname{Diff}(\mathbb{S}^1), \quad V \mapsto \operatorname{Fl}_1^V$$

assigning to a (time-independent) vector field its time 1-flow.

Recall that $\theta\colon \mathbb{R} \to \mathbb{R}/2\pi \cong \mathbb{S}^1$, $\theta \mapsto e^{i\theta}$ is a submersion. Composing the submersion with a vector field of the circle, we identify vector fields of the

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

circle with 2π-periodic maps $\mathbb{R} \to \mathbb{R}$. For a constant vector field $X^c(\theta) \equiv c$, a quick computation shows that its flow, $\exp(tX^c(\theta)) = e^{i(\theta+tc)}$, is for fixed t a rotation of the circle.

Step 1: A diffeomorphism $\eta \in \mathrm{Diff}(\mathbb{S}^1)$ without fixed points is an exponential of $V \in \mathcal{V}(\mathbb{S}^1)$ if and only if η is conjugate to a rotation.
If $V \in \mathcal{V}(\mathbb{S}^1)$ has a zero, its exponential has a fixed point. Thus if $\eta = \exp(X)$, we must have $X(\theta) \neq 0$. We will now construct $\varphi \in \mathrm{Diff}(\mathbb{S}^1)$ and a constant vector field X^c such that $\varphi \circ \eta = \exp(X^c) \circ \varphi$. Assume for a moment that for all $t \in [0,1]$ we have the identity $\varphi \circ \exp(tX) = \exp(tX^c) \circ \varphi$. Differentiating at $t = 0$ yields with Lemma 3.40 the identity $T_\theta\varphi(X(\theta)) = X^c(\varphi(\theta))$ for all $\theta \in \mathbb{S}^1$. For ease of computation identify now φ and X with periodic mappings $\mathbb{R} \to \mathbb{R}$. Then the equation reads $\varphi'(\theta)X(\theta) = c$ for all $\theta \in \mathbb{R}$. Integrating, we obtain, with $\varphi(0) = 0$ and the fact that X vanishes nowhere, that $\varphi(\theta) = \int_0^\theta c/X(s)\mathrm{d}s$. We will now choose c such that $\varphi(\theta + 2\pi) - \varphi(\theta) = \int_t^{t+2\pi} c/X(s)\mathrm{d}s = 2\pi$ (because then φ descends to a diffeomorphism of $\mathbb{S}^1$). Since X is 2π-periodic, we can simply choose $c = 2\pi \int_0^{2\pi} X(s)\mathrm{d}s$. Since the flows of the φ-related vector fields X and $T\varphi^{-1} \circ X \circ \varphi$ are conjugate by φ, we obtain $\varphi \circ \eta = \varphi \circ \exp(X) = \exp(X^c) \circ \varphi = R_c \circ \varphi$. Thus if a fixed-point-free diffeomorphism of $\mathbb{S}^1$ is the exponential of a vector field, it is conjugate to a rotation.

Step 2: Diffeomorphisms near the identity which are not exponentials.
We claim that there are diffeomorphisms φ, arbitrarily near the identity $\mathrm{id}_{\mathbb{S}^1}$, such that:

(a) φ has no fixed points;
(b) there exists $\theta_0 \in \mathbb{S}^1$ and $1 < n \in \mathbb{N}$ such that $\varphi^n(\theta_0) = \theta_0$, but $\varphi^n \neq \mathrm{id}_{\mathbb{S}^1}$.

If this were true, then we note that if φ is an exponential, (a) and Step 1 imply that it must be conjugate to rotation. However, this is impossible by (b), as a rotation which has a point of period n must itself be periodic with period n. Thus the image of the exponential does not contain φ and $\mathrm{Diff}(\mathbb{S}^1)$ cannot be locally exponential.

To construct a diffeomorphism with properties (a) and (b) consider for $n \in \mathbb{N}$ large enough and $0 < \varepsilon < 1/n$ the maps

$$f_{n,\varepsilon}(\theta) = \theta + \frac{\pi}{n} + \varepsilon \sin^2(n\theta)$$

descend to diffeomorphisms of the circle which satisfy (a)–(b) and can be made arbitrarily close to the identity (i.e. every identity neighbourhood in the compact open C^∞-topology contains an $f_{n,\varepsilon}$ for n large enough). We leave the details as Exercise 3.3.6.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

More generally one can prove that the diffeomorphism group of a compact manifold is not locally exponential.

3.43 Example (a) As a consequence of Lemma 3.40 and the inverse function theorem, every Banach–Lie group and thus, in particular, every finite-dimensional Lie group is locally exponential.

(b) The unit group of a Mackey-complete CIA A (see Example 3.35) is locally exponential (Glöckner and Neeb, 2012).

(c) If G is a locally exponential Lie group, we shall see in Exercise 3.4.6 that the current groups $C^\infty(K,G)$ (see §3.4) are locally exponential. Thanks to Wockel (2007) this generalises even to the group of gauge transformations of a principal bundle with locally exponential gauge group.

3.44 Remark While in infinite dimensions not every closed subgroup of a (locally exponential) Lie group is again a Lie group (see Example 3.11), there are conditions which ensure that a closed subgroup of a locally exponential Lie group is a Lie subgroup. We refer to Neeb (2006, IV) for more information.

Exercises

3.3.1 A smooth map $\alpha\colon \mathbb{R} \to G$ to a Lie group is called a 1-*parameter subgroup* of (a regular Lie group) G if it is a group homomorphism, that is, $\alpha(s+t) = \alpha(s)\alpha(t)$ for all $s,t \in \mathbb{R}$.

(a) Show that a 1-parameter subgroup α is an integral curve of the left-invariant vector field L_v and the right-invariant field R_v, generated by $v := \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \alpha(t)$, i.e. $\dot\alpha(t) = L_v(\alpha(t))$ and $\dot\alpha(t) = R_v(\alpha(t))$.

(b) Show that if $\alpha,\beta\colon \mathbb{R} \to G$ are smooth, $\alpha(0) = \mathbf{1} = \beta(0)$ with $\dot\alpha(t) = L_v(\alpha(t))$ and $\dot\beta(t) = R_v(\beta(t))$ (the associated right-invariant vector field) then $\alpha = \beta$ and these curves are a 1-parameter subgroup.
Hint: To show equality differentiate $\alpha(t)\beta(-t)$; for the second statement take the derivative of $\beta(t-s)\beta(s)$.

3.3.2 Consider the multiplicative group $(\mathbb{R}^\times,\cdot)$ as a Lie group. Show that $\delta^\ell(f)(t) = f'(t)/f(t)$ and explain the name logarithmic derivative.

3.3.3 Let $\gamma_i\colon [0,1] \to G, i = 1,2$ be smooth curves. Show that $\delta^\ell(\gamma_1) = \delta^\ell(\gamma_2)$ if and only if $\gamma_1(t) = g\gamma_2(t)$, for all $t \in [0,1]$ and some fixed $g \in G$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

3.3.4 Consider a locally convex space $(E,+)$ as a Lie group. Show that the Lie bracket on $\mathbf{L}(E) = E$ vanishes and the Lie group exponential is $\exp_E = \mathrm{id}_E$.

3.3.5 Let $\alpha\colon G \to H$ be a Lie group morphism between regular Lie groups. Show that:

(a) for $\gamma \in C^1([0,1],G)$ we have $\delta^\ell(\alpha\circ\gamma) = \mathbf{L}(\alpha)(\delta^\ell(\gamma))$;

(b) for $\eta \in C^\infty([0,1],\mathbf{L}(G))$ one has $\mathrm{Evol}(\mathbf{L}(\alpha)\circ\eta) = \alpha\circ\mathrm{Evol}(\eta)$ (a similar formula holds for evol). Further, prove the naturality of the Lie group exponential, that is,

$$\exp_H \circ \mathbf{L}(\alpha) = \alpha \circ \exp_G . \tag{3.8}$$

3.3.6 Consider for $n \in \mathbb{N}$ and $0 < \varepsilon < 1/n$ the maps $f_{n,\varepsilon}\colon \mathbb{R} \to \mathbb{R}$, $\theta \mapsto \theta + \pi/n + \varepsilon \sin^2(n\theta)$.

(a) Show that $f_{n,\varepsilon}$ descends via the submersion $\mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z}$ to a diffeomorphism $\varphi_{n,\varepsilon}$ of $\mathbb{S}^1$. Moreover, prove that there are $\varphi_{n,\varepsilon}$ arbitrarily near the identity (where $\mathrm{Diff}(\mathbb{S}^1) \subseteq C^\infty(\mathbb{S}^1,\mathbb{S}^2)$ carries the compact open C^∞-topology).
Hint: Control $f_{n,\varepsilon}$ in the compact open C^∞-topology on $\mathbb{R}$.

(b) Show that $\varphi_{n,\varepsilon}$ does not possess a fixed point, but since $f_{n,\varepsilon}^{2n}(0) = 0$ (modulo 2π) it has a periodic point of period $2n$.

(c) Show that the $2n$-periodic orbit $\varphi_{n,\varepsilon}^k(0)$, $k = 1,\dots,2n-1$ is unique, that is, if θ is not contained in the orbit, then $\varphi_{n,\varepsilon}^{2n}(\theta) \neq \theta$. We deduce that $\varphi_{n,\varepsilon}^{2n} \neq \mathrm{id}_{\mathbb{S}^1}$.

3.3.7 Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. By Lie's third theorem (Hilgert and Neeb, 2012, Theorem 9.4.11) there exists a connected, simply connected Lie group G such that $\mathbf{L}(G) = \mathfrak{g}$. Assume that M is a compact manifold such that $\phi\colon \mathfrak{g} \to \mathcal{V}(M)$ is a Lie algebra morphism (with respect to the negative of the usual bracket of vector fields). Prove that ϕ induces a smooth action of G on M (this is a very restricted version of the Lie–Palais theorem (Palais, 1957)).

3.4 The Current Groups

In this section, we construct a class of infinite-dimensional Lie groups which occur naturally in theoretical physics: loop groups and current groups. As a first step in the construction we prove a useful local characterisation of Lie groups.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Notation (Multiplicative Notation for Sets) In the statement of Proposition 3.45 we use multiplicative notation for sets. Recall that $A \cdot B$ means the set of all elements which can be written as a product of elements in A and B (in that order).

3.45 Proposition (Bourbaki, 1998, Ch. III, §1, No. 9 Proposition 18) *Let G be a group and $U, V \subseteq G$ such that $\mathbf{1} \in V = V^{-1} := \{g \in G \mid g^{-1} \in V\}$ and $V \cdot V \subseteq U$. Assume that U is equipped with a (smooth) manifold structure such that $V \subseteq U$ and the mappings $\iota|_V^V : V \to V$ and $m_G|_{V\times V} : V \times V \to U$ are smooth. Then the following holds:*

(a) *There is a unique manifold structure on the subgroup*

$$G_0 := \langle V \rangle := \{v_1 \cdots v_k \mid v_i \in V, k \in \mathbb{N}\}$$

such that G_0 becomes a Lie group, $V \subseteq G_0$, and G_0 and U induce the same manifold structure on V.

(b) *Assume that for each g in a generating set of G, there is $\mathbf{1} \in W_g \subseteq U$ such that $gW_g g^{-1} \subseteq U$ and $c_g : W_g \to U$, $h \mapsto ghg^{-1}$ is smooth. Then there is a unique manifold structure on G turning G into a Lie group such that $V \subseteq G$ and both G and U induce the same manifold structure on V.*

Proof *Step 1: Shifted open sets*. If $A \subseteq V$ and $v_0 \in V$ with $v_0 A \subseteq V$, then $v_0 A$ is open in V as the preimage of A under the smooth map $\delta_{v_0} : V \to U$, $g \mapsto v_0^{-1} g$.

Step 2: Shifted charts. Pick $W \subseteq V$ such that $W = W^{-1}$, $W^3 = W \cdot W \cdot W \subseteq U$ and there exists a manifold chart (φ, W) of U. For $g \in G$ we consider the map $\varphi_g : gW \to \varphi(W)$, $h \mapsto \varphi(g^{-1}h)$. The idea is now to construct an atlas from these shifted charts which is compatible with the manifold structure on V.

Step 3: Manifold induced by shifted charts. Observe that for $g_1 W \cap g_2 W \neq \emptyset$ we have $g_2^{-1} g_1 \in W^2$ (this entails $g_1^{-1} g_2 \in W^2$). Then Step 1 yields $W \cap g_2^{-1} g_1 W \subseteq W$, whence $D_{g_1,g_2} := \varphi_{g_1}(g_1 W \cap g_2 W) = \varphi(W \cap g_2^{-1} g_1 W)$ is open in the model space. For $d \in D_{g_1,g_2}$ we have now

$$\varphi_{g_2} \circ \varphi_{g_1}^{-1}(d) = \varphi(g_2^{-1} g_1 \varphi^{-1}(d)),$$

which is a smooth mapping by choice of W, and so the change of charts is smooth. Endow G with the final topology with respect to the atlas $\mathcal{S} := (\varphi_g)_{g\in G}$ (this topology is Hausdorff since W is Hausdorff). We conclude that $\mathcal{S}$ is a manifold atlas turning G into a smooth manifold.

Moreover, for each $g_0 \in G$ we have $\varphi_{g g_0} \circ \lambda_{g_0}|_{gW} = \varphi_g$, whence the left translation with elements in G is smooth for this manifold structure.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Step 4: The manifold structures coincide on V. By our assumptions on multiplication and inversion, we can find for every $v_0 \in V$ a set $\mathbf{1} \in A \subseteq W$ such that $v_0 A \subseteq V$. Step 3 shows that $v_0 A$ is open in G and $A \to v_0 A$, $h \mapsto v_0 h$ is a diffeomorphism with respect to the structure induced by $\mathcal{S}$: in particular, $V \subseteq G$. Now Step 1 implies that this holds for the structure induced by U, whence both coincide on V.

We are now in a position to prove both claims of the proposition:

(a) By definition we have $G_0 = \langle V \rangle = \bigcup_{k \in \mathbb{N}} V^k$. We can write $V^2 = \bigcup_{g \in V} gV \subseteq G$ and thus, inductively, $V^k \subseteq G$ and $G_0 \subseteq G$. Denote by $(G, \mathcal{S})$ the manifold induced by the atlas $\mathcal{S}$. Then Step 4 shows that $\delta \colon V \times V \to G_0 \subseteq (G, \mathcal{S})$, $(g,h) \mapsto gh^{-1}$ is smooth and this is equivalent to the smoothness of multiplication and inversion on the set V. For $g_0, h_0, g, h \in G$ we let $c_{h_0}(a) := h_0 a h_0^{-1}$ and obtain the identity

$$\begin{aligned} \delta(g,h) &= (h_0 g_0^{-1})^{-1} h_0 (g_0^{-1} g)(h_0^{-1} h)^{-1} h_0^{-1} \\ &= \lambda_{g_0^{-1} h_0} \circ c_{h_0} \circ \delta(\lambda_{g_0^{-1}}(g), \lambda_{h_0^{-1}}(h)). \end{aligned} \tag{3.9}$$

We proceed now by induction on k and show first that δ is smooth on $V^k \times V$. For $k = 1$ we know that δ is smooth. However, note that also $c_{h_0} \colon V \to G_0$ is smooth since

$$c_{h_0}(g) = \lambda_{h_0} \circ \delta(g, h_0) \tag{3.10}$$

and δ is smooth on $V \times V$. Moreover, if $g \in V^k$ we can pick $g_0 \in V^{k-1}$ such that for all x in a g-neighbourhood $g_0^{-1} x \in V$. Following again (3.10), we see that

$$c_{h_0}(x) = \lambda_{h_0 g_0} \delta(g_0^{-1} x, h_0) \text{ is smooth for all } h_0 \in H \text{ and } x \text{ near } g.$$

We conclude that $c_{h_0} \colon G_0 \to G_0$ is smooth for each $h_0 \in V$.

Thus for the induction step from k to $k+1$ we see now that $\delta \colon V^{k+1} \times V \to G_0$ is given in a neighbourhood of (g,h) by the composition of smooth maps (3.9) for $h_0 := h$ and $g_0 \in V$ with $g_0^{-1} g \in V^k$. It follows that $\delta \colon G_0 \times V \to G_0$ is smooth. Now a trivial induction together with (3.9) shows that $\delta \colon G_0 \times G_0 \to G_0$ is smooth. We conclude that G_0 is a Lie group.

(b) We prove inductively that δ is smooth of an element (g,h) where g is arbitrary and $h = s_1 \cdots s_n$ for s_i, $i = 1, \ldots, n$, in the generating set $\mathcal{G}$. Starting with $n = 1$ we choose $h_0 := s_1^{-1}$ and $g_0 := g^{-1}$. By (a) the map δ is smooth on G_0, whence there is an open (g,h)-neighbourhood which gets mapped by $\delta \circ (\lambda_{g_0}, \lambda_{h_0})$ into W_g. Since $h_0 = s_1 \in \mathcal{G}$ by the assumption, (3.9) implies that δ is smooth in a neighbourhood of (g,h). Now by (3.10), conjugation with $h_0 \in \mathcal{G}$ is smooth in a g-neighbourhood, hence on all

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

of G. If δ is smooth on an open neighbourhood of $G \times \bigcup_{1\leq k\leq n} \mathcal{G}^k$ and $h = s_1 \cdots s_{n+1}$, we argue as above: Taking $h_0 = s_1^{-1}$ shows that δ is smooth in a neighbourhood of $G \times \bigcup_{1\leq k\leq n+1} \mathcal{G}^k$ for all $n \in \mathbb{N}$. As smoothness of δ is equivalent to the smoothness of the group operations, we deduce that G is a Lie group.

The uniqueness of the manifold structure of the Lie groups G_0 and G in (a) and (b) follows from the fact that any other manifold structure with these properties induces the same manifold structure on the open $\mathbf{1}$-neighbourhood V. □

We will now use Proposition 3.45 to construct the Lie group structure for the current group $C^\infty(K,G)$. The crucial insight here is that for a Lie group, an atlas of charts can be constructed by left translating a chart at the identity. We will see in the construction of the current group that this allows us to construct a canonical manifold of mappings by left translating the pushforward of a suitable chart at the identity.

General Assumption For the rest of this section we let K be a compact manifold, G a (possibly infinite-dimensional) Lie group with Lie algebra $\mathbf{L}(G)$. Then we choose and fix $\mathbf{1} \in U_1 \Subset G$ together with a chart $\varphi\colon U_1 \to U \subseteq \mathbf{L}(G)$ such that $\varphi(\mathbf{1}) = 0$ and $T_{\mathbf{1}}\varphi = \mathrm{id}_{\mathbf{L}(G)}$. Further, we pick $\mathbf{1} \in V_1 \Subset U_1$ such that $V_1 \cdot V_1 \subseteq U_1, V_1^{-1} = V_1$ and set $V := \varphi(V_1)$. Mapping spaces will as always be endowed with the compact open C^∞-topology.

Let us now exploit again the local formulation of the Lie group structure of G in a chart around the identity (as in the identification of the Lie algebra, 3.21). In order to clearly distinguish the local group operations from the (globally defined) group operations and the pushforwards appearing on the function spaces, we introduce new symbols for the local operations.

3.46 The mappings

$$\mu\colon V \times V \to U, \mu(x,y) := \varphi(\varphi^{-1}(x)\varphi^{-1}(y)), \quad \iota\colon V \to V, \iota(x) := \varphi(\varphi^{-1}(x)^{-1})$$

are smooth. Moreover, $C^\infty(K,U_1)$ is an open subset of $C^\infty(K,G)$ and we equip it with the smooth manifold structure, making the bijection $C^\infty(M,U_1) \to C^\infty(K,U)$, $\gamma \mapsto \varphi \circ \gamma$ a diffeomorphism of smooth manifolds. Note that thanks to Exercise 2.3.2, the manifold $C^\infty(K,U)$ is canonical. Then the pushforwards $\mu_*\colon C^\infty(K,V\times V) = C^\infty(K,V)\times C^\infty(K,V) \to C^\infty(K,U)$ and $\iota_*\colon C^\infty(K,V) \to C^\infty(K,V)$ are smooth by Corollary 2.19.

3.47 Theorem *The pointwise operations turn $C^\infty(K,G)$ into a Lie group, a* current group. *Its Lie algebra is $C^\infty(K,\mathbf{L}(G))$ with the pointwise Lie bracket (a* current algebra*).*

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Proof We check that the assumptions of Proposition 3.45 are satisfied: The identity element in $C^\infty(K,G)$ is the constant map $\mathbf{1}_{C^\infty(K,G)}(k) = \mathbf{1}_G$. Thus 3.46 shows that $C^\infty(K,U) \subseteq C^\infty(K,G)$ is an open identity neighbourhood such that multiplication and inversion are smooth on the smaller neighbourhood $C^\infty(K,V)$. Now let $\gamma \in C^\infty(K,G)$; by compactness of $\gamma(K) \subseteq G$ there are $\mathbf{1} \in W_1 \subseteq V_1$ and $\gamma(K) \subseteq P \subseteq G$ such that $pW_1p^{-1} \subseteq U_1$ for all $p \in P$. Set $W := \varphi(W_1)$ and note that since $C^\infty(K,W) \subseteq C^\infty(K,V)$ so is $C^\infty(K,W_1) \subseteq C^\infty(K,V_1)$. Since $c\colon P \times W_1 \to U_1, h(p,w) := pwp^{-1}$ is smooth, so is

$$h_\gamma = \varphi \circ h \circ (\gamma \times \varphi^{-1})\colon K \times W \to U, h_\gamma(x,y) = \varphi(\gamma(x)\varphi^{-1}(y)\gamma(x)^{-1}).$$

Since the manifolds $C^\infty(K,W)$ and $C^\infty(K,U)$ are canonical (see 3.46), Proposition 2.17 yields a smooth map $(h_\gamma)_\star\colon C^\infty(K,W) \to C^\infty(K,U)$, $\eta \mapsto h_\gamma \circ (\mathrm{id}_W \times \eta)$. Now conjugation by γ coincides on $C^\infty(K,W)$ with $(h_\gamma)_\star$ and is thus smooth on $C^\infty(K,W) \subseteq C^\infty(K,G)$. Now Proposition 3.45 provides a unique smooth Lie group structure for $C^\infty(K,G)$.

To identify the Lie algebra, note that $\varphi_*\colon C^\infty(K,U_1) \to C^\infty(K,U)$, $\gamma \mapsto \varphi \circ \gamma$ is a chart around the identity of the Lie group $C^\infty(K,G)$. Exploiting that $T_\mathbf{1}\varphi = \mathrm{id}_{\mathbf{L}(G)}$, we use 2.22 to identify $T_{\mathbf{1}_{C^\infty(K,G)}}\varphi_* = (T_\mathbf{1}\varphi)_* = (\mathrm{id}_{\mathbf{L}(G)})_*$ whence $\mathbf{L}(C^\infty(K,G) \cong C^\infty(K,\mathbf{L}(G))$. Now the point evaluations $\mathrm{ev}_x\colon C^\infty(K,G) \to G$, $\gamma \mapsto \gamma(x)$ are Lie group morphisms (Exercise 3.4.2), and so $\mathbf{L}(\mathrm{ev}_x) = \mathrm{ev}_x\colon C^\infty(K,\mathbf{L}(G))$ is a Lie algebra morphism by Lemma 3.27. This implies $[\gamma,\eta](x) = [\gamma(x),\eta(x)]$ and thus the bracket is given by the pointwise bracket. □

3.48 Remark For non-compact manifolds M, the group $C^\infty(M,G)$ can, in general, not be made a Lie group. To see this, consider $\mathbb{N}$ a 0-dimensional manifold. Then $C^\infty(\mathbb{N},\mathbb{S}^1) \cong (\mathbb{S}^1)^\mathbb{N} := \prod_{n\in\mathbb{N}} \mathbb{S}^1$ is a compact topological group. However, since it is not locally contractible (see Exercise 3.4.5) it cannot be a manifold and thus it cannot be a Lie group.

We shall now prove that regularity of the target Lie group G is inherited by the current group $C^\infty(K,G)$. The idea is simple: The Lie type equation on the current group is just a parametric version of the Lie-type equation on the target Lie group. In other words, if $\eta\colon [0,1] \to C^\infty(K,\mathbf{L}(G))$ is a smooth curve, then we obtain for each $k \in K$ a Lie type equation

$$\frac{\partial}{\partial t}\gamma^\wedge(t,k) = (\dot\gamma(t))(k) = (\gamma(t).\eta(t))(k) = \gamma^\wedge(t,k).\eta^\wedge(t,k).$$

Thanks to the regularity of the target Lie group G, we can solve the Lie type equation (uniquely) for each fixed k. Then the solutions to these equations

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

glue back together to a solution on the current group (because of a suitable exponential law).

3.49 Proposition *If G is a regular Lie group, then $C^\infty(K,G)$ is a regular Lie group.*

Proof In view of the preliminary considerations, we see that the pointwise solutions to the Lie type equation must be the solution to the equation on the current group. It remains to apply the exponential law[5] to show that these solutions depend smoothly on the initial data. We obtain a canonical isomorphism Θ:

$$C^\infty([0,1],C^\infty(K,\mathbf{L}(G))\cong C^\infty([0,1]\times K,\mathbf{L}(G))\cong C^\infty(K,C^\infty([0,1],\mathbf{L}(G)).$$

The map $h := (\operatorname{evol}_G)_*\circ\Theta\colon C^\infty([0,1],C^\infty(K,\mathbf{L}(G))\to C^\infty(K,G)$ is smooth by Corollary 2.19 as $\operatorname{evol}_G\colon C^\infty([0,1],\mathbf{L}(G))\to G$ is smooth. Let $\operatorname{ev}_x\colon C^\infty(K,G)\to G$ and $e_x\colon C^\infty(K,\mathbf{L}(G))\to\mathbf{L}(G)$ be the point evaluations in x. We note that $\mathbf{L}(\operatorname{ev}_x)=e_x$. Hence Exercise 3.3.5 yields $\operatorname{ev}_x\circ h=\operatorname{evol}_G\circ(e_x)_*\circ\Theta$ for all $x\in K$. Since the evaluations separate the points on $C^\infty(K,G)$, this implies $h=\operatorname{evol}_{C^\infty(K,G)}$. □

Similar to regularity being hereditary, current groups inherit the property of being locally exponential from their target Lie group. We leave this as Exercise 3.4.6. In the following sections, we discuss two special cases of current groups and their subgroups: loop groups and groups of gauge transformations (associated to a principal bundle).

Loop Groups

If $K=\mathbb{S}^1$ the current group $\mathrm{LG}:=C^\infty(\mathbb{S}^1,G)$ is better known as a *loop group* (see Pressley and Segal, 1986). Much is known about loop groups and their representation theory. We mention that they are, in particular, connected to the representation theory of Kac–Moody Lie algebras and to quantum field theory; see Schmid (2010).

3.50 Remark In the literature the group $C(\mathbb{S}^1,G)$ of continuous loops is also often called the loop group LG of G. For us the loop group will nevertheless consist always of smooth loops.

[5] Since [0, 1] has boundary, we cannot use the exponential law, Theorem 2.1.2. Instead, we have to appeal to the stronger version (Alzaareer and Schmeding, 2015, Theorem A). Note that the result is the same and, in particular, the (partial) derivative of the adjoint map corresponds to the derivative.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

We shall discuss two canonical subgroups of LG for finite-dimensional G.

3.51 The canonical group morphism $I\colon G \to C^\infty(\mathbb{S}^1,G)$, $g \mapsto (k \mapsto g)$ identifies G with the subgroup of constant loops. As $C^\infty(\mathbb{S}^1,G)$ is a canonical manifold by Exercise 3.4.3, smoothness of I is equivalent to smoothness of $I^\wedge\colon G\times\mathbb{S}^1 \to G$, $I^\wedge(g,k) = g$ (which is obvious). Further, a local argument in charts around $(x,k) \in G\times\mathbb{S}^1$ shows that the partial derivative of $I^\wedge$ identifies the derivative of I. Thus by C.12 we obtain

$$TI(v_g) = (k \mapsto v_g \in T_gG) \in C^\infty(\mathbb{S}^1,TG) \cong TC^\infty(\mathbb{S}^1,G), \quad g \in G,\ v_g \in T_gG$$

and this map is clearly injective for every g. Thus I is infinitesimally injective and since G is finite dimensional, it is an immersion. Moreover, I is a Lie group morphism with smooth inverse: Thanks to Exercise 3.4.2(c), the evaluation maps $\mathrm{ev}_k : LG \to G$, $k \in \mathbb{S}^1$ are Lie group morphisms. Choosing, for example, $h := \mathrm{ev}_{\mathbf{1}_{\mathbb{S}^1}}\big|_{I(G)}$, we obviously have $h \circ I = \mathrm{id}_G$, whence I is a topological embedding onto its image. Hence I is a smooth embedding and if we identify G with $I(G)$, we can think of G as a Lie subgroup of LG (see Glöckner, 2016, Lemma 1.13).

Now define the group of all loops starting at $\mathbf{1}_G$ as $\Omega G := \{f \in LG \mid f(\mathbf{1}_{\mathbb{S}^1}) = \mathbf{1}_G\}$. Since $\mathrm{ev}_{\mathbf{1}_{\mathbb{S}^1}} : C^\infty(\mathbb{S}^1,G) \to G$ is a Lie group morphism and a submersion with $\Omega G = \mathrm{ev}^{-1}_{\mathbf{1}_{\mathbb{S}^1}}(\mathbf{1}_G)$, we see that ΩG is a split submanifold of LG, hence a Lie subgroup. Summarising,

$$\mathbf{1} \longrightarrow \Omega G \hookrightarrow \mathrm{LG} \underset{I}{\overset{\mathrm{ev}_{\mathbf{1}_{\mathbb{S}^1}}}{\rightleftarrows}} G \longrightarrow \mathbf{1}$$

is a split sequence of Lie groups, whence $\mathrm{LG} \cong \Omega G \rtimes G$. We finally note that this allows one to define the *fundamental homogeneous space*

$$\mathrm{LG}/G = \{[h] \mid f \in [h] \text{ if } f = g\cdot h \text{ for some } g \in G\} \ (\cong \Omega G),$$

which plays an important role in the theory of loop groups (see Pressley and Segal, 1986).

Groups of Gauge Transformations

We briefly discuss an important class of infinite-dimensional Lie groups also arising in physics. In extremely broad strokes the setup is as follows: Fix a manifold M. In physics this would correspond to spacetime. Physical fields in the spacetime M are described as sections of certain bundles $P \to M$ called principal bundles (see Definition 3.52). A premier example here is the electromagnetic field. In this formulation, the famous Maxwell equations – see

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Example 3.54 – can be interpreted as differential equations for a certain connection on the principal bundle. We will not describe connections on principal bundles here, but point out that the geometric data of the connection corresponds to the potential of the field. However, the symmetry group of the space of connections on the bundle is the so-called group of gauge transformations. We shall describe it as an infinite-dimensional Lie group and identify it as a subgroup of the current groups studied earlier. Let us begin by recalling the notion of a principal bundle. Much more information on principal bundles can be found in the usual finite-dimensional literature (Baum, 2014; Husemoller, 1993).

3.52 Definition Let G be a Lie group with a smooth right action $\rho\colon E\times G\to E$. Assume that the quotient $p\colon E\to M:=E/G$ is a smooth locally trivial fibre bundle with typical fibre F, that is, p is a smooth map such that there is an open cover $(U_i)_{i\in I}$ of M and ρ-equivariant diffeomorphisms (*bundle trivialisations*) $\kappa_i\colon p^{-1}(U_i)\to U\times F$ conjugating p to the projection. The quadruple (E,p,M,F) is a *principal G-bundle* if the action ρ is simply transitive, that is, for each $f_0\in F\cong p^{-1}(x)$, we have $G\to F$, $g\mapsto f_0 g$ is a diffeomorphism.

The group G in the definition of the principal bundle is also called the *structure group* of the principal bundle. In the physics literature it is customary to call the structure group the *gauge group* of the principal bundle. We will follow the physics terminology and warn the reader that in the literature, the term 'gauge group' is also used for the group of gauge transformations we are about to define.

3.53 Remark Note that for a principal G-bundle the fibre F is diffeomorphic to the gauge group G, but since F lacks a preferred choice of unit element, there is no preferred group structure on the fibre F (one also says that F is a *G-torsor*[6]).

3.54 Example If M is a smooth manifold and G a Lie group, then the trivial bundle $\mathrm{pr}_M\colon M\times G\to M$ is a principal bundle.

While on first sight nothing interesting is happening here, trivial principal bundles appear in interesting physics applications. As an example, we would like to mention Maxwell's equations from electromagnetics. Fix a contractible $U\subseteq\mathbb{R}^3$. The electrical field E, the magnetic field H and the current J in U can be described by time-dependent vector fields on U. Denote by ρ the electrical charge and by c the speed of light. Then Maxwell's equations can be expressed in differential operator notation as

[6] See https://math.ucr.edu/home/baez/torsors.html for more examples and explanations.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$\operatorname{curl} E = -\frac{1}{c}\frac{\partial H}{\partial t}, \quad \operatorname{div} H = 0, \quad \operatorname{curl} H = \frac{1}{c}\frac{\partial E}{\partial t} + \frac{4\pi}{c}J, \quad \operatorname{div} E = 4\pi\rho.$$

Now one can show that the Maxwell equations can be interpreted as differential equations for the curvature of a $\mathbb{S}^1$-connection on the trivial $\mathbb{S}^1$-bundle $(U\times\mathbb{R})\times \mathbb{S}^1 \to U\times\mathbb{R}$. We refer to Baum (2014, Section 7.1) for the derivation and more information.

Interpreting the Maxwell equations in the language of principal bundles turns out to be a fruitful idea. Replacing the trivial bundle by a bundle with a (non-abelian) structure group, one arrives at gauge theories such as the Yang–Mills theory or the Chern–Simons theory.

3.55 Example Let K be a compact Lie group and G a closed Lie subgroup of K; then one can show that the quotient $M := K/G$ is again a manifold such that the canonical mapping $q\colon K \to K/G = M$ becomes a submersion (this is a so-called *homogeneous space*). Then (K,q,M,G) is a principal G-bundle (Bröcker and Tom Dieck, 1995, Chapter I, Theorem 4.3).

As a more concrete example, consider the compact Lie group $\mathrm{SO}_3(\mathbb{R})$ (see Example 3.3(b)). Recall from Duistermaat and Kolk (2000, 1.2.A) that $\mathrm{SO}_3(\mathbb{R})$ can be identified with the 3-dimensional unit sphere $\mathbb{S}^3$ in $\mathbb{R}^4$. Moreover, the spherical group $\mathbb{S}^1$ embeds into $\mathrm{SO}_3(\mathbb{R})$ as the rotation subgroup $\left\{\begin{bmatrix}\cos(\alpha) & -\sin(\alpha) & 0\\ \sin(\alpha) & \cos(\alpha) & 0\\ 0 & 0 & 1\end{bmatrix}, \alpha\in\mathbb{R}\right\}$. This identification allows us to interpret multiplication with $\mathbb{S}^1$-elements as rotations on $\mathbb{S}^3$ and it is not hard to see that the quotient, $\mathbb{S}^3/\mathbb{S}^1 = \mathrm{SO}_3(\mathbb{R})/\mathbb{S}^1$ is diffeomorphic to the 2-dimensional sphere $\mathbb{S}^2$. The resulting homogeneous space

$$\mathbb{S}^1 \hookrightarrow \mathbb{S}^3 \xrightarrow{\ q\ } \mathbb{S}^2$$

is known in differential topology as the Hopf fibration; see Sharpe (1997, Example 3.14). Note that the Hopf fibration is not isomorphic (as an $\mathbb{S}^1$-principal bundle) to the trivial $\mathbb{S}^1$-bundle over $\mathbb{S}^2$ (Duistermaat and Kolk, 2000, p. 56).

3.56 Definition Let (E,p,M,F) be a principal G-bundle. Then define the *group of gauge transformations* of the bundle as follows:

$$\mathrm{Gau}(E) := \{\varphi\in \mathrm{Diff}(E) \mid p\circ\varphi = p, \text{ for all } g\in G,\ \varphi\circ\rho(\cdot,g) = \rho(\cdot,g)\circ\varphi\}.$$

3.57 If φ is a gauge transformation we can identify $\varphi(e) = \rho(e,f(e))$ for some smooth function $f\colon E\to G$, and the relation $\varphi\circ\rho(\cdot,g) = \rho(\cdot,g)\circ\varphi$ then yields

$$f(\rho(e,g)) = g^{-1}f(e)g, \quad \text{for all } e\in E,\ g\in G. \tag{3.11}$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Conversely, every smooth function $f\colon E \to G$ satisfying (3.11) defines a gauge transformation via $\varphi_f(e) \coloneqq \rho(e, f(e))$. We will show in Exercise 3.4.7 that for

$$C^\infty(E,G)^G \coloneqq \{f \in C^\infty(E,G) \mid \text{for all } e \in E,\ g \in G, f(\rho(e,g)) = g^{-1} f(e) g\},$$

the map

$$C^\infty(E,G)^G \to \mathrm{Gau}(E), \quad f \mapsto \varphi_f, \quad \varphi_f(e) = \rho(e, f(e)) \tag{3.12}$$

is a group isomorphism. Hence the group of gauge transformations can naturally be identified as a subgroup of the current group.

3.58 Example If $\mathrm{pr}_M\colon M \times G \to M$ is a trivial principal G-bundle, it is easy to see that $\mathrm{Gau}(M \times G) \cong C^\infty(M,G)$. So if M is compact, the group of gauge transformations inherits a Lie group structure from the current group in this case.

For the trivial $\mathbb{S}^1$-principal bundle connected to Maxwell's equations, Example 3.54, the group of gauge transformations (aka the current group in this case) acts on $\mathbb{S}^1$-connections on the trivial bundle (connections in the same orbit are called gauge equivalent). As already mentioned, the Maxwell equations can be interpreted as a differential equation for the curvature of a $\mathbb{S}^1$-connection. See Baum (2014, Section 7.1).

3.59 Remark If G is a locally exponential Lie group and M a compact manifold, then the group of gauge transformations $\mathrm{Gau}(E)$ of a smooth principal G-bundle (E,p,M,F) carries a natural Lie group structure, turning it into a locally exponential Lie group (Wockel, 2006). Indeed one can show that the group of gauge transformations then becomes a Lie subgroup of a finite product of suitable current groups (Wockel, 2007).

Alternatively, the Lie group structure of the group of gauge transformations can be derived by identifying the group of gauge transformations with the vertical bisections of a gauge groupoid; see Remark 6.17 and Example 6.13.

Exercises

3.4.1 We establish some auxiliary results for the proof of Proposition 3.45.

(a) Let G be a group which is also a manifold. Show that G is a Lie group if and only if the mapping $\delta\colon G \times G \to G$, $(g,h) \mapsto gh^{-1}$ is smooth.

(b) Let G be a Lie group (actually a topological group would be enough for the following) and $V \subseteq G$ with $\mathbf{1} \in V$. Show that there is $\mathbf{1} \in W \subseteq V$ with $W^{-1} = W$ and $W \cdot W \subseteq V$. (Can you generalise this to $W^n \subseteq V, n \in \mathbb{N}$?)

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(c) Let G, H be Lie groups and $f\colon G \to H$ be a group homomorphism. Prove that f is a Lie group morphism if and only if there exists an open $\mathbf{1}_G$-neighbourhood U such that $f|_U$ is smooth.

(d) Let G be a group with two manifold structures $\mathcal{S}$ and $\mathcal{T}$ turning G into a Lie group. Show that $(G,\mathcal{S})$ and $(G,\mathcal{T})$ are diffeomorphic (as manifolds and then also as Lie groups) if and only if there is an open $\mathbf{1}$-neighbourhood on which the two manifold structures coincide.

3.4.2 Verify several details from the proof of Theorem 3.47:

(a) Let $K \subseteq G$ be a compact subset of a Lie group G and $\mathbf{1} \in U \subseteq G$. Show that there is an open $\mathbf{1}$-neighbourhood W and an open K-neighbourhood P such that $pWp^{-1} \subseteq U$ for all $p \in P$.

(b) Verify that the final topology, with respect to the charts $(\varphi_g)_{g\in G}$ from Step 3 of the proof, is Hausdorff.

(c) Show that the point evaluations $\mathrm{ev}_x\colon C^\infty(K,G) \to G, \gamma \mapsto \gamma(x)$ are Lie group morphisms.
Hint: For group homomorphisms it suffices to check smoothness in an identity neighbourhood.

3.4.3 Prove that the current group $C^\infty(K,G)$ is a canonical manifold.
Hint: For a Lie group, it suffices to construct an open identity neighbourhood whose manifold structure is canonical.

3.4.4 By C.2 every Lie group G admits a local addition. Show that the manifold structure on $C^\infty(K,G)$ constructed in C.10 coincides with the one from Theorem 3.47.

3.4.5 Show that $C^\infty(\mathbb{N},\mathbb{S}^1) = \prod_{n\in\mathbb{N}} \mathbb{S}^1$ is not locally contractible, that is, show that if N is a neighbourhood of a point $x \in \prod_{n\in\mathbb{N}} \mathbb{S}^1$ then there is no continuous mapping $h\colon [0,1]\times N \to N$ such that $h(0,\cdot) = \mathrm{id}_N$ and $h(0,y) = x$, for all $y \in N$.

3.4.6 Let G be a locally exponential Lie group. Show that then also $C^\infty(K,G)$ is locally exponential.
Hint: Study Proposition 3.49 and note, in addition, that $\exp_G$ induces a chart for G.

3.4.7 Show that the map $C^\infty(E,G)^G \to \mathrm{Gau}(E)$, $f \mapsto \varphi_f$ from 3.57 is a group isomorphism.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

4

Lifting Geometry to Mapping Spaces II: (Weak) Riemannian Metrics

In this chapter, we will discuss Riemannian metrics on infinite-dimensional spaces. Particular emphasis will be placed on the new challenges which arise on infinite-dimensional spaces.

General Assumption for This Chapter To ensure the existence of integrals we shall always assume that the manifolds in this chapter are modelled on (Mackey) complete locally convex spaces.

4.1 Weak and Strong Riemannian Metrics

Riemannian metrics come in several flavours on infinite-dimensional spaces which are not present in the finite-dimensional setting. The strongest flavour (as we shall see) is the notion of a strong Riemannian metric which is treated in classical monographs such as Lang (1999) and Klingenberg (1995).

4.1 Definition Let M be a manifold modelled on a locally convex space E. A *weak Riemannian metric* g on M is a smooth map

$$g : TM \oplus TM \to \mathbb{R}, \qquad (v_x, w_x) \mapsto g_x(v_x, w_x)$$

(where $TM \oplus TM$ is the Whitney sum, see §C.9) satisfying

(a) $g_x := g|_{T_xM \times T_xM}$ is symmetric and bilinear for all $x \in M$,
(b) $g_x(v,v) \geq 0$ for all $v \in T_xM$ with $g_x(v,v) = 0$ if and only if $v = 0$.

If a weak Riemannian metric satisfies in addition, for all $x \in M$,

(c) the topology of the inner product space (T_xM, g_x) coincides with the topology of T_xM as a subspace of TM,

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

we say that g is a *strong Riemannian metric*. A manifold with a weak (/strong) Riemannian metric will be called a *weak (/strong) Riemannian manifold*.

4.2 Example Every Hilbert space $(H, \langle\cdot,\cdot\rangle)$ becomes a strong Riemannian manifold using the identifications $TH = H \times H$, $TH \oplus TH = H^3$ and setting $g_c(v,w) := \langle v, w\rangle$.

4.3 Example Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space. The locally convex space $C^\infty([0,1],H)$ with the compact-open C^∞-topology is a Fréchet space (but not a Banach or Hilbert space!). Consider the L^2-inner product on this space as

$$\langle f,g\rangle_{L^2} := \int_0^1 \langle f(t), g(t)\rangle \mathrm{d}t.$$

This is a bilinear map on $C^\infty([0,1],H)$. By construction of the compact open C^∞-topology, the inclusion of $C^\infty([0,1],H) \to C([0,1],H)_{\text{c.o.}}$ is continuous linear and it is not hard to prove that the mapping $\beta\colon C([0,1],H)^2 \to H$, $(f,g) \mapsto \int_0^1 \langle f(t), g(t)\rangle \mathrm{d}t$ is continuous bilinear. In conclusion, the L^2-inner product is a continuous bilinear form on $C^\infty([0,1],H)$. Interpreting the locally convex space as a manifold, we have

$$\begin{aligned} TC^\infty([0,1],H) &= C^\infty([0,1],H) \times C^\infty([0,1],H) \cong C^\infty([0,1], H\times H) \\ &= C^\infty([0,1], TH). \end{aligned}$$

We obtain an isomorphism $TC^\infty([0,1],H) \oplus TC^\infty([0,1],H) \cong C^\infty([0,1],H^3)$ (cf. also §C.13) which transforms $f_c, g_c \in T_c C^\infty([0,1],H)$ into a triple (c,f,g) of H-valued functions. Thus

$$g_{L^2}\colon TC^\infty([0,1],H) \oplus TC^\infty([0,1],H) \cong C^\infty([0,1],H^3) \to \mathbb{R}, \quad (c,f,g) \mapsto \langle f,g\rangle_{L^2}$$

is a weak Riemannian metric, called the *L^2-metric*. Note that the L^2-metric is not a strong Riemannian metric as the topology of the inner product space is the one induced by the inclusion $C^\infty([0,1],H) \subseteq L^2([0,1],H)$ where the space on the right-hand side is the space of all square-integrable functions with the L^2-topology.

4.4 If (M,g) is a weak Riemannian manifold, there is an injective linear map

$$\begin{aligned} \flat\colon TM &\to T^*M = \bigcup_{x\in M} L(T_xM, \mathbb{R}), \\ T_xM \ni v &\mapsto g_x(v,\cdot), \end{aligned}$$

where T^*M is the dual tangent bundle; see Remark 1.45. If M is a finite-dimensional manifold, then it follows by counting dimensions that $\flat$ is (fibre-wise) an isomorphism and thus $\flat$ is an isomorphism of vector bundles;

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Similarly, if g is a strong Riemannian metric, $\flat$ is an isomorphism of vector bundles; see Proposition 4.5. The inverse of $\flat$ in these situations is often denoted by $\sharp$ and the pair of isomorphisms is known as the musical isomorphisms of a Riemannian manifold. For infinite-dimensional manifolds, the map $\flat$ will, in general, not be surjective.

The next result yields a useful characterisation of strong Riemannian metrics (which is the definition of a strong Riemannian metric in the classical texts such as Lang (1999) and Klingenberg (1995)). We just mention that for the proof some tools from functional analysis are required (which we cite but we do not provide a detailed review here).

4.5 Proposition *Let (M,g) be a weak Riemannian manifold. The following conditions are equivalent:*

(a) *g is a strong Riemannian metric;*
(b) *M is a Hilbert manifold and $\flat\colon TM \to T^*M$ is surjective;*
(c) *M is a Hilbert manifold and $\flat\colon TM \to T^*M$ is a vector bundle isomorphism.*

In particular, on every finite-dimensional manifold M, a weak Riemannian metric is automatically strong.

Proof *Step 1: Strong Riemannian metric implies Hilbert and surjective $\flat$.* If g is a strong Riemannian metric, every tangent space T_xM is a Hilbert space (since it is Mackey-complete; see Remark 1.13). As the tangent spaces are isomorphic to the modelling spaces of M, we see that M is a Hilbert manifold. Now the surjectivity of $\flat$ is a consequence of the Riesz representation theorem (Meise and Vogt, 1997, 11.9).

Step 2: $\flat$ surjective on a Hilbert manifold implies that $\flat$ is a bundle isomorphism. Note first that since M is a Hilbert manifold, there is a canonical smooth structure on the dual bundle (see Remark 1.45), whence it makes sense to consider $\flat$ as a smooth map between TM and T^*M. By construction, we know that $\flat$ induces continuous vector space isomorphisms $T_xM \to (T_xM)^*$ for every tangent space. Hence the open mapping theorem (Rudin, 1991, I. 2.11) shows that $\flat$ is fibre-wise an isomorphism, which together with Lang (1999, III. Proposition 1.3) shows that $\flat$ is a vector bundle isomorphism.

Step 3: Hilbert manifold and $\flat$ being a bundle isomorphism imply that g is strong. Consider $x \in M$ and note that as g is continuous on $TM \oplus TM$, g_x is continuous with respect to the Hilbert space topology of the fibre. Let us denote the Hilbert space norm on T_xM by $\|\cdot\|$. Then the ball $B_1^{g_x}(0) := \{y \in T_xM \mid g_x(y,y) < 1\}$ is by construction $B_1^{g_x}(0)$ a disc 0-neighbourhood

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(see Definition A.6) and the norm $\|\cdot\|_{g_x}$ induced by g_x (aka the Minkowski functional of the ball, Lemma A.16) is continuous. As the unit ball in $(T_xM, \|\cdot\|)$ is bounded, this shows already that there is a constant $K > 0$ with $\|\cdot\| \leq K\|\cdot\|_{g_x}$. To see that the norms are equivalent, we have to prove that $B^1_{g_x}(0)$ is bounded in the Hilbert space topology. As $\flat$ is surjective, every bounded linear functional on T_xM is of the form $\flat(v)$ for some $v \in T_xM$. Applying the Cauchy–Schwarz inequality, we derive

$$\sup_{y \in B_1^{g_x}(0)} |\flat(v)(y)| = \sup_{y \in B_1^{g_x}(0)} |g_x(v,y)| \leq \|v\|_{g_x}.$$

We conclude that every bounded linear functional is bounded on $B_1^{g_x}(0)$. This property is called weakly bounded and it is known that on a Hilbert space, every weakly bounded set in a locally convex space is bounded (Rudin, 1991, Theorem 3.13). Summing up, g_x induces the Hilbert space topology of T_xM and since x was arbitrary, this shows that g is a strong Riemannian metric. □

4.6 Example The space of immersions $\mathrm{Imm}(\mathbb{S}^1, \mathbb{R}^d)$ is an open subset of $C^\infty(\mathbb{S}^1, \mathbb{R}^d)$ by Lemma 2.6. We endow it with a weak Riemannian metric (an invariant version of the L^2-metric, §4.2)

$$g_c(h,k) := \int_{\mathbb{S}^1} \langle h(\theta), k(\theta)\rangle \|\dot{c}\| \mathrm{d}\theta.$$

Recall that since $\mathrm{Imm}(\mathbb{S}^1, \mathbb{R}^d) \subseteq C^\infty(\mathbb{S}^1, \mathbb{R}^d)$, we have an identification

$$T_c \,\mathrm{Imm}(\mathbb{S}^1, \mathbb{R}^d) = C^\infty(\mathbb{S}^1, \mathbb{R}^d);$$

see C.12. Working in the identification we immediately see that the musical morphism reduces to the map

$$\flat_c(h) = \|\dot{c}\| \cdot h, h \in C^\infty(\mathbb{S}^1, \mathbb{R}^d),$$

where the dot signifies pointwise multiplication. So the image of $T_c \,\mathrm{Imm}(\mathbb{S}^1, \mathbb{R}^d)$ under $\flat_c$ can be identified as $C^\infty(\mathbb{S}^1, \mathbb{R}^d)$. However, the (topological) dual space of $C^\infty(\mathbb{S}^1, \mathbb{R}^d)$ is the larger space $\mathcal{D}'(\mathbb{S}^1)^d$ of $\mathbb{R}^d$-valued distributions (see Taylor, 2011, Section 3).

4.7 Example Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space with $S_H = \{x \in H \mid \|x\| = 1\}$ (Example 1.34). The Hilbert sphere is a strong Riemannian manifold with the induced metric $g_x(v,w) := \langle v, w\rangle$ (where $T_xS_H = \{v \in H \mid \langle v, x\rangle = 0\}$).

The distinction between strong and weak Riemannian metrics has far-reaching consequences (see e.g. §4.2 on geodesic distance).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Exercises

4.1.1 Verify Example 4.2: Every Hilbert space $(H, \langle\cdot,\cdot\rangle)$ is a strong Riemannian manifold.

4.1.2 Show that the mapping $\mu\colon C([0,1],H)^2 \to \mathbb{R}$, $(f,g) \mapsto \langle f,g\rangle_{L^2}$ is continuous bilinear (hence smooth), see Example 4.3.
Hint: $\int_0^1 : C([0,1],\mathbb{R}) \to \mathbb{R}$, $f \mapsto \int_0^1 f(t)\mathrm{d}t$ is linear. Exploit that $C([0,1],\mathbb{R})$ is a Banach space to prove continuity via usual integral estimates.

4.1.3 Verify that the Hilbert sphere is a strong Riemannian manifold (Example 4.7).

4.1.4 We shall treat Riemannian metrics on spaces of smooth functions on the sphere.
Hint: If you are not familiar with integration on manifolds: Use $\theta\colon [0,2\pi] \to \mathbb{S}^1$, $t \mapsto (\cos(t),\sin(t))$ to reduce the integral to a usual integral; see note at the beginning of Chapter 5 and compare Lee (2013, Chapter 16).

(a) Let (M,g) be a strong Riemannian metric. Use C.13 to show that the L^2-metric

$$g_c^{L^2}(h,k) := \int_{\mathbb{S}^1} g_{c(\theta)}(h(\theta),k(\theta))\mathrm{d}\theta$$

defines a weak Riemannian metric on $C^\infty(\mathbb{S}^1,M)$.

(b) Verify that the metric in Example 4.6 is a weak Riemannian metric.

4.2 The Geodesic Distance on a Riemannian Manifold

General Assumption In this section (M,g) denote a *strong* Riemannian manifold if nothing else is said. For convenience, we shall always assume that M is connected.

Having a Riemannian metric at our disposal, we can define the length of curves.

4.8 Definition Let $c\colon [a,b] \to M$ be a piecewise C^1-curve.[1] Then we define the *length* and the *energy* of c as

[1] That is, there exists a partition of $[a, b]$ into subintervals such that on each of them the curve restricts to a C^1-curve.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$\mathrm{Len}(c) := \int_a^b \sqrt{g_{c(t)}(\dot{c}(t),\dot{c}(t))}\mathrm{d}t,$$
$$\mathrm{En}(c) := \frac{1}{2}\int_a^b g_{c(t)}(\dot{c}(t),\dot{c}(t))\mathrm{d}t.$$

For $x, y \in M$ then define

$$\Gamma(x,y) := \{c\colon [0,1] \to M \mid c(0) = x, c(1) = y, \text{ and } c \text{ is piecewise } C^1\}.$$

Finally, we define the *geodesic distance* between points $x, y \in M$ as

$$\mathrm{dist}(x,y) := \inf_{c\in\Gamma(x,y)} \mathrm{Len}(c) = \inf_{c\in\Gamma(x,y)} \int_0^1 \sqrt{g_{c(t)}(\dot{c}(t),\dot{c}(t))}\mathrm{d}t.$$

Due to the Cauchy–Schwarz inequality, for curves $c\colon [a,b] \to M$ we find (see Klingenberg, 1995, Proposition 1.8.7) that

$$\mathrm{Len}(c)^2 \le 2\mathrm{En}(c)(b-a) \qquad \text{(with equality if and only if } \dot{c} \text{ is constant).} \tag{4.1}$$

4.9 Remark Note that since every interval $[a,b]$ is diffeomorphic to $[0,1]$, the chain rule implies that the definition of $\Gamma(x,y)$ and of the geodesic distance does not depend on $[0,1]$. It is only a convenient choice for us and we shall ignore this choice in the construction of paths to avoid cumbersome reparametrisation arguments.

4.10 Lemma *The geodesic distance is a pseudo-distance, that is,*

(a) $\mathrm{dist}(x,y) \ge 0$ *for* $x, y \in M$,
(b) $\mathrm{dist}(x,y) = \mathrm{dist}(y,x)$ *for all* $x, y \in M$,
(c) $\mathrm{dist}(x,z) \le \mathrm{dist}(x,y) + \mathrm{dist}(y,z)$ *for all* $x, y, z \in M$.

The proof of this lemma is left as an exercise.

On strong Riemannian manifolds, the geodesic distance is also *point separating*, that is, in addition to the properties from Lemma 4.10, it satisfies

(d) $\mathrm{dist}(x,y) \ne 0$ for all $x, y \in M$ with $x \ne y$.

Moreover, one can prove the following result for strong Riemannian metrics.

4.11 Theorem (Klingenberg, 1995, Theorem 1.9.5) *Let* (M,g) *be a strong Riemannian metric. The function* dist$\colon M \times M \to \mathbb{R}$ *defines a metric on* M. *The topology derived from* dist *coincides with the given topology of* M.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

However, in general, *on infinite-dimensional manifolds* the geodesic distance might not be a metric. Indeed it might be non-point separating and even stronger, the geodesic distance might be vanishing.[2] Note that if c is a path connecting $x \neq y$, then $\mathrm{Len}(c) > 0$ (or in other words $\mathrm{Len}(c) = 0$ implies that the path is constant). Thus the vanishing of $\mathrm{dist}(x, y)$ means that we can find arbitrarily short paths connecting the two points. We showcase this in the following example.

4.12 Example (Magnani and Tiberio, 2020) Consider the space $(\ell^2, \langle\cdot,\cdot\rangle)$ of all square-summable real sequences; see Example 3.11. The map $A\colon \ell^2 \to \ell^2$, $(x_n)_{n\in\mathbb{N}} \mapsto (\frac{1}{n^3}x_n)_{n\in\mathbb{N}}$ is continuous linear and induces a bilinear symmetric map $B\colon \ell^2 \times \ell^2 \to \mathbb{R}, B(\mathbf{x},\mathbf{y}) \coloneqq \langle \mathbf{x}, A\mathbf{y}\rangle$. Identifying the tangent spaces of ℓ^2, we obtain a weak Riemannian metric via

$$g\colon T\ell^2 \oplus T\ell^2 = \bigcup_{\mathbf{p}\in\ell^2} \ell^2 \times \ell^2 \to \mathbb{R}, \quad (T_{\mathbf{p}}\ell^2)^2 \ni (\mathbf{x},\mathbf{y}) \mapsto e^{-\|\mathbf{p}\|^2} B(\mathbf{x},\mathbf{y}).$$

We will now prove that the weak Riemannian manifold (ℓ^2, g) has vanishing geodesic distance, that is, that for every $\mathbf{p} \neq \mathbf{q}$ in ℓ^2, we have $\mathrm{dist}(\mathbf{p},\mathbf{q}) = 0$. To this end, let $\mathbf{e}_n$ be the sequence with 1 in the nth place and zeroes everywhere else. We construct a path from $\mathbf{p}$ to $\mathbf{q}$ via

$$c_n\colon [0,1] \to \ell^2, \quad t \mapsto \begin{cases} \mathbf{p} + 3tn\mathbf{e}_n, & t \in [0,1/3], \\ \mathbf{p} + n\mathbf{e}_n + (3t-1)(\mathbf{q}-\mathbf{p}), & t \in [1/3,2/3], \\ \mathbf{q} + (3-3t)n\mathbf{e}_n, & t \in [2/3,1]. \end{cases}$$

By construction, c_n is a piecewise linear curve connecting $\mathbf{p}$ to $\mathbf{q}$ and passing through $\mathbf{p}+n\mathbf{e}_n$ and $\mathbf{q}+n\mathbf{e}_n$ on the way. We claim that $\mathrm{Len}(c_n) \to 0$ as $n \to \infty$ and thus $\mathrm{dist}(\mathbf{p},\mathbf{q}) = 0$. To see this, we observe that

$$c_n'(t) = 3n\mathbf{e}_n, t \in [0,1/3[, \quad c_n'(t) = 3(\mathbf{q}-\mathbf{p}), t \in]1/3,2/3[,$$
$$c_n'(t) = -3n\mathbf{e}_n,\ t \in]2/3,1].$$

Moreover, since $\mathbf{p},\mathbf{q} \in \ell^2$ there is $N > 0$ such that every component of $\mathbf{p}$ and $\mathbf{q}$ with $n > N$ satisfies $|\pi_n(\mathbf{p})|, |\pi_n(\mathbf{q})| < 1/3$ (here π_n is the projection on the nth component). Hence we see that for every such n we have $\pi_n(\mathbf{p}+t(\mathbf{q}-\mathbf{p})) > -1$.

[2] We say the geodesic distance is *non-vanishing* if there exist $x, y \in M$ such that $\mathrm{dist}(x, y) \neq 0$. So every point separating geodesic distance is non-vanishing but not vice versa.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

We now estimate the length $\mathrm{Len}(c_n)$:

$$\begin{aligned}\int_0^1 c_n(t)\mathrm{d}t &= 3\left(\int_0^{1/3} e^{-\|c_n(t)\|^2}\sqrt{B(n\mathbf{e}_n,n\mathbf{e}_n)}\mathrm{d}t\right.\\ &\quad+\int_{1/3}^{2/3} e^{-\|c_n(t)\|^2}\underbrace{\sqrt{B(\mathbf{q}-\mathbf{p},\mathbf{q}-\mathbf{p})}}_{=:K>0}\mathrm{d}t\\ &\quad\left.+\int_{2/3}^1 e^{-\|c_n(t)\|^2}\sqrt{B(-n\mathbf{e}_n,-n\mathbf{e}_n)}\mathrm{d}t\right)\\ &\le 3\left(\int_0^{1/3}\sqrt{\langle n\mathbf{e}_n,\frac{1}{n^2}\mathbf{e}_n\rangle}\mathrm{d}t+K\int_{1/3}^{2/3} e^{-(n^2+\|\mathbf{p}+(3t-1)(\mathbf{q}-\mathbf{p})\|^2+2n\langle\mathbf{e}_n,\mathbf{p}+(3t-1)(\mathbf{q}-\mathbf{p})\rangle)}\mathrm{d}t\right.\\ &\quad\left.+\int_{2/3}^1\sqrt{B(-n\mathbf{e}_n,-n\mathbf{e}_n)}\mathrm{d}t\right)\\ &\le \frac{1}{\sqrt{n}}+3K\int_{1/3}^{2/3} e^{-n^2-2n\pi_n(\mathbf{p}+t(\mathbf{q}-\mathbf{p}))}\mathrm{d}t+\frac{1}{\sqrt{n}}\overset{n>N}{\le}\frac{2}{\sqrt{n}}+Ke^{-n^2+2n}.\end{aligned}$$

We conclude that the length of the curve c_n converges to zero as $n\to\infty$ whence the geodesic distance vanishes.

Another interesting example in this regard is the invariant L^2-metric on the group of circle diffeomorphisms. Before we state this example, let us exhibit a general construction principle for (weak) Riemannian metrics on Lie groups.

4.13 (Invariant metrics on a Lie group) Let G be a (perhaps infinite-dimensional) Lie group with Lie algebra $\mathbf{L}(G)$. Assume that $\langle\cdot,\cdot\rangle\colon \mathbf{L}(G)\times\mathbf{L}(G)\to\mathbb{R}$ is a continuous inner product on the Lie algebra. Then we define a *right invariant (weak) Riemannian metric* via the following formula:

$$\langle V,W\rangle_g := \langle T\rho_g^{-1}(V),T\rho_g^{-1}(W)\rangle,\quad \text{for all } V,W\in T_gG. \tag{4.2}$$

Here ρ_g is the right translation by g and we remark that due to the smoothness of the group operations the resulting metric is indeed a (weak) Riemannian metric. By construction, the right invariant metric is invariant under the right action of the Lie group G on TG via right multiplication.

Note that by replacing every ρ_g in (4.2) by the left translation λ_g, we can obtain a *left-invariant (weak) Riemannian metric* associated to the given inner product.

4.14 Example (Right-invariant L^2-metric on $\mathrm{Diff}(\mathbb{S}^1)$) We consider $\mathrm{Diff}(\mathbb{S}^1)$ again as an open subset of $C^\infty(\mathbb{S}^1,\mathbb{S}^1)$. Recall from Example 3.5 that this

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

manifold structure turns the diffeomorphism group into a Lie group. Moreover, $T_\varphi \operatorname{Diff}(\mathbb{S}^1) = C^\infty_\varphi(\mathbb{S}^1, T\mathbb{S}^1) \cong C^\infty_\varphi(\mathbb{S}^1, \mathbb{S}^1 \times \mathbb{R}) \cong C^\infty(\mathbb{S}^1, \mathbb{R})$.[3]

On $C^\infty(\mathbb{S}^1, \mathbb{R})$ we have the inner product (where we refer to the discussion in the beginning of Chapter 5 for the meaning of the integral):

$$\langle u, v\rangle_{L^2} := \int_{\mathbb{S}^1} f(\theta) g(\theta) \mathrm{d}\theta.$$

Plugging this into (4.2), we obtain the *(right)-invariant L^2-metric*

$$g^{L^2,\mathrm{inv}}_\varphi(u, v) := \langle u \circ \varphi^{-1}, v \circ \varphi^{-1}\rangle_{L^2}.$$

Thanks to a theorem of Michor and Mumford, the geodesic distance with respect to this metric vanishes; see, for example, Kolev (2008, Theorem 4).

Exercises

4.2.1 Establish the estimate (4.1).

4.2.2 Prove Lemma 4.10 and verify that the pseudodistance does not depend on the choice of interval (Remark 4.9).

4.2.3 The following provide the details for Example 4.12:

(a) Show that the map A makes sense and is linear and continuous and thus the map B is bilinear and symmetric.

(b) Prove that g is a weak Riemannian metric on ℓ^2.

4.2.4 Prove that the construction of right- (or left-)invariant metrics on a Lie group from 4.13 yields a weak Riemannian metric.
Hint: To check smoothness of the metric, use Lemma 3.12 to identify the Whitney sum $TG \oplus TG$.

Geodesics on Infinite-dimensional Manifolds (Informal Discussion)

To get a better understanding of the distance on weak Riemannian metrics, it seems useful to study curves 'of shortest length' between two points the so-called *geodesics*. Before we study geodesics in the next section, we discuss some aspects in an informal way.

If the manifold is a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ viewed as a strong Riemannian manifold (Example 4.2), the curve of shortest length between $a, b \in H$ is the straight line

$$\mathbb{R} \ni t \mapsto \gamma(t) := (b - a)t + a \in H. \tag{4.3}$$

[3] Here we exploit that $\mathbb{S}^1$ is a Lie group, whence the tangent bundle $T\mathbb{S}^1 \cong \mathbb{S}^1 \times \mathbb{R}$ is trivial.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Note that this curve also satisfies $\operatorname{dist}(\gamma(t),\gamma(s)) = \|b-a\||t-s|$, and so it realises the shortest distance between any two given points on the line. On a manifold, we would like to compute curves which satisfy the same property at least locally, that is, in a neighbourhood of each point the curve realises the shortest distance for every pair of points on the curve.[4] Thanks to (4.1), one can equivalently describe geodesics between p and q as curves of minimal energy, that is, extrema of the energy En restricted to $\Gamma(p,q)$. Hence to find geodesics we consider the derivative of the energy. Working locally in a chart (U,φ) of M (and suppressing most identifications in the notation; see Lemma C.17), this yields, for the derivative $d\mathrm{En}(c;h)$, the formula

$$\int_0^1 \frac{1}{2} d_1 g_U(c,c'(t),c'(t);h)) - d_1 g(c(t),h(t),c'(t);c'(t))) \\ - g_U(c(t),h(t),c''(t))\mathrm{d}t.$$

To find the geodesics, one would now have to isolate h in the expression to rewrite the differential in the form $\int_0^1 g_c(\ldots,h)\mathrm{d}t$ and extract the geodesic equation. In general, this is not possible, as the $\flat$-map is not an isomorphism, whence the existence of geodesics for weak Riemannian metrics is a priori unclear (see also Proposition 4.5).

We remark that the geodesics of weak Riemannian metrics are also of independent interest in the context of Euler–Arnold theory (see Chapter 7). There, certain partial differential equations can be interpreted as geodesic equations on infinite-dimensional manifolds.

4.15 Example (Inviscid Burgers equation) One can show (Kolev, 2008, 3.2) that geodesics with respect to the invariant L^2-metric from Example 4.14 correspond to solutions of the inviscid Burgers equation (also known as the Hopf equation):

$$u_t + 3uu_x = 0 \text{ (subscripts denoting partial derivatives).}$$

We shall investigate a similar situation later in Example 7.4.

The observant reader should have noted that we are talking about geodesics (which are the solutions of the inviscid Burgers equation) for a weak Riemannian metric which was described in Example 4.14 as having vanishing geodesic distance. It might be tempting to think that this implies that all geodesics must be constant (since only these curves have length 0 and geodesics were

[4] The sphere $\mathbb{S}^1$ is a Riemannian manifold by Example 4.7. For $x \in \mathbb{S}^1$ consider a closed curve running in a great circle from x around the sphere. Then this curve realises the shortest distance from x to another point y until it passes the point antipodal to x. However, as long as we restrict to an open neighbourhood which does not contain points antipodal to each other, the curve realises the shortest distance from one point to the other.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

described as smooth curves (locally) minimising the length between their endpoints). This, however, is wrong (the reader might either consult the partial differential equations (PDE) literature, or observe – see Example 7.4 – that the related equation $u_t + uu_x = 0$ is the geodesic equation of a weak Riemannian metric with non-vanishing geodesic distance). This equation is the Burgers equation and it admits non-constant solutions. So what's wrong here? For strong Riemannian metrics one can show that every geodesic is (locally) length minimising. However, there are also other characterisations of geodesics (which we will discuss in the next section) which coincide with our informal definition for strong Riemannian metrics. In the weak setting, the situation is (as always) more complicated.

Exercises

4.2.5 We verify some details concerning geodesics in Hilbert space.

(a) Show that the path $c(t) = (b-a)t + a, t \in [0,1]$, (4.3), realises the shortest length with respect to all C^1-curves connecting a to b in the Hilbert space H.
Hint: Prove first for a path consisting of two line segments meeting at a point p that its length is longer than the length of c if p does not lie on c. Then use that C^1-paths are rectifiable.

(b) Prove the distance property claimed after (4.3): For all $s,t \in \mathbb{R}$, $\mathrm{dist}(\gamma(s),\gamma(t)) = \|b-a\||t-s|$. Note that this is (up to reparametrisation to unit speed) the geodesic property for curves in metric spaces; see Bridson and Haefliger (1999, Definition 1.3).

4.3 Geodesics, Sprays and Covariant Derivatives

In this section, we consider three objects associated to a Riemannian metric (see e.g. Lang, 1999; Abraham et al., 1988). This will enable us to study geodesics on Riemannian manifolds. The idea is that every (strong) Riemannian metric induces a covariant derivative, a metric spray and a connector. These can be used to conveniently describe geodesics. The main point is to introduce the concepts of spray and covariant derivative while we relegate many details of the constructions to the literature, for example, Lang (1999, Chapter VIII).

General Assumption In this section, (M,g) will denote a *strong* Riemannian manifold if nothing else is said. Thus there will be no need to worry about the existence of solutions to certain differential equations. Note that for weak

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Riemannian metrics similar computations are in principle possible, but require the careful checking of several technical details.[5]

We first define certain vector fields on the tangent bundle $\pi_M : TM \to M$ which are called *sprays*. For any spray, geodesics can be defined and if the spray is the metric spray associated to a Riemannian metric, these geodesics will turn out to be the geodesics of the Riemannian metric.

4.16 Definition (Spray) A vector field $S \in \mathcal{V}(TM)$ on the tangent manifold TM is said to be of *second order* if

$$T\pi_M(S(v)) = v \text{ for all } v \in TM.$$

For each $t \in \mathbb{R}$ we denote by $t_{TM} : TM \to TM$ the vector bundle morphism which is given in each fibre by multiplication with t. Now a second-order vector field S is a *spray* if

$$S(tv) = T(t_{TM})(t \cdot S(v)) \text{ for all } t \in \mathbb{R}, v \in TM. \tag{4.4}$$

To understand the meaning of (4.4), let us localise in a chart. (Warning: The next identities hold up only to identifications in charts, which we will suppress in the notation!)

4.17 Choose $U \subseteq M$ such that $TU \cong U \times E$ (where E is the model space of M). Then $TTU \cong (U \times E) \times (E \times E)$. Now if S is a second-order vector field, its restriction to U is given for $(u,V) \in U \times E$ by $S(u,V) = ((u,V),V,S_2(u,V))$. If S is a spray, the equation (4.4) reads on U as follows:

$$S(u,tV) = (u,tV,tV,S_{U,2}(u,tV)) = (u,tV,tV,t^2 S_{U,2}(u,V)).$$

The map $S_{U,2}$ is thus quadratic with respect to scalar multiplication in the fibre. Furthermore, this implies together with Exercise 4.3.1 that $S_{U,2}(x,v) = \frac{1}{2} d_2^2 S_2((x,0);(v,v))$.

We define $B\colon M \to C^\infty(TM \oplus TM, TM)$ as the map which is locally on a chart domain (U,φ) given by the symmetric bilinear map $B_U(x;v,w) := d_2^2 S_{U,2}((x,0);(v,w))$. Associated to this bilinear map is the quadratic form $Q_U(x,v) := B_U(x;v,v)$.[6] Note that in finite-dimensional Riemannian geometry $-B$ is represented by the so-called *Christoffel symbols* (Lang, 1999, pp. 213–214).

4.18 (Integral curves of second-order vector fields) Let S be a smooth second-order vector field on the manifold M and assume that S admits integral curves

[5] We shall do this in Chapter 5 for the L^2-metric on $C^\infty(\mathbb{S}^1, M)$.

[6] We recall that due to the identity $B_U(x;v,w) = \frac{1}{2}(Q_U(x, v+w) - Q_U(x,v) - Q_U(x,w))$, the quadratic form carries exactly the same information as the bilinear form.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(i.e. C^1-curves $\beta\colon J \to TM$ with $S(\beta) = \dot{\beta}$[7]). Note that an integral curve $\beta\colon J \to TM$ of a second-order vector field S satisfies $(\pi_M \circ \beta)\dot{}(t) = \beta(t)$.

4.19 Definition A C^2-curve $\alpha\colon J \to M$ is a *geodesic* of the spray S if $\dot{\alpha}\colon J \to TM$ is an integral curve of S, or equivalently, if α satisfies the *geodesic equation*

$$\ddot{\alpha}(t) = S\left(\dot{\alpha}(t)\right), \quad \text{for all } t \in J.$$

Note that $\dot{\alpha}(t) = T_t\alpha(1) \in TM$. Equivalently, the geodesic equation becomes (in local coordinates) the equation $\ddot{\alpha}(t) = B_{\alpha(t)}(\dot{\alpha}(t), \dot{\alpha}(t))$.

4.20 Example A strong Riemannian metric g induces a spray $S^g\colon TM \to T^2M$ which we describe locally on a chart domain $U \subseteq M, TU = U \times E, T^2U = (U \times E) \times E \times E$ (again suppressing the identifications!). Namely, we think of the local representative $g_U : U \times E \times E \to \mathbb{R}$, $(x, v, w) \mapsto g_x(v, w)$ of the metric as a mapping with three components. Then we define the associated spray via $S^g_U(x, v) := ((x, v), (v, \Gamma_U(x, v)))$ for $v \in T_xM$, where the quadratic form Γ is the unique map which satisfies for all $v, w \in T_xU$.

$$g_U(x, \Gamma_U(x, v), w) = \frac{1}{2} d_1 g_U(x, v, v; w) - d_1 g_U(x, v, w; v), \tag{4.5}$$

Here $d_1 g_U$ denotes the partial derivative with respect to the first component (see Proposition 1.20). We leave the verification that the local mappings S^g_U yield a spray S^g on TM to the reader (Exercise 4.3.2). Note that the formula (4.5) makes sense for weak Riemannian metrics, but since $\flat$ is, in general, not surjective for a weak Riemannian metric (see Proposition 4.5), it is a priori not clear whether the condition defines a unique map Γ_U.

4.21 Definition (Metric spray) Let (M, g) be a weak Riemannian manifold. A spray S is called *metric spray* if its associated quadratic form Q satisfies for each chart (U, φ) the formula (4.5), that is, for all $v, w \in T_xU$.

$$g_U(x, Q_U(x, v), w) = \frac{1}{2} d_1 g_U(x, v, v; w) - d_1 g_U(x, v, w; v), \tag{4.6}$$

Let us point out that the metric spray S^g associated to a Riemannian metric g can be interpreted in the sense of Hamiltonian mechanics: It can be shown (see Abraham et al., 1988, p. 493) that the spray S^g is the Hamiltonian vector field on TM associated with the kinetic energy function

[7] If the manifold M is modelled on a Banach space, such curves always exist due to the standard ODE solution theory. Beyond Banach spaces, the existence of such curves needs to be established for the special case at hand.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$e\colon TM \to \mathbb{R}$, $v_x \mapsto \frac{1}{2}g_x(v_x, v_x)$. This has several interesting consequences, such as conservation of energy by the geodesics (again we refer to Abraham et al., 1988, Supplement 8.1.B, for more information). We shall return to the relation between energy and geodesics in Chapter 7.

The next example shows that there is not necessarily a metric spray. It was pointed out to me by C. Maor and we urge the reader to compare it with Theorem 4.30.

4.22 Example (A weak Riemannian metric without metric spray) We return to the weak Riemannian metric on the Hilbert space $(\ell^2.\langle\cdot,\cdot\rangle)$ from Example 4.12:

$$g\colon T\ell^2 \oplus T\ell^2 = \bigcup_{\mathbf{p}\in\ell^2} \ell^2 \times \ell^2 \to \mathbb{R}, \quad T_{\mathbf{p}}\ell^2 \ni (\mathbf{x}, \quad \mathbf{y}) \mapsto e^{-\|\mathbf{p}\|^2} \sum_{n=1}^{\infty} \frac{\mathbf{x}_n \mathbf{y}_n}{n^3},$$

where the subscripts denote elements of a sequence. A quick computation shows that

$$d_1 g(\mathbf{p}, \mathbf{x}, \mathbf{y}; \mathbf{w}) = -2\langle \mathbf{p}, \mathbf{w}\rangle g(\mathbf{p}, \mathbf{x}, \mathbf{y}) = -2g(\mathbf{p}, (n^3\mathbf{p}_n)_{n\in\mathbb{N}}, \mathbf{w}) \sum_{n=1}^{\infty} \frac{\mathbf{x}_n \mathbf{y}_n}{n^3},$$

where the last equality only makes sense if the sequence $(n^3\mathbf{p}_n)_{n\in\mathbb{N}}$ is contained in ℓ^2. Plugging this identity into the right-hand side of (4.6), we see that

$$\begin{aligned}
&\frac{1}{2} d_1 g(\mathbf{p}, \mathbf{x}, \mathbf{x}; \mathbf{w}) - d_1 g(\mathbf{p}, \mathbf{x}, \mathbf{w}; \mathbf{x}) \\
&\quad = -g(\mathbf{p}, (n^3\mathbf{p}_n)_{n\in\mathbb{N}}, \mathbf{w}) \sum_{n=1}^{\infty} \frac{\mathbf{x}_n^2}{n^2} + 2\langle \mathbf{p}, \mathbf{x}\rangle g(\mathbf{p}, \mathbf{x}, \mathbf{w}) \\
&\quad = g\Big(\mathbf{p}, 2\langle \mathbf{p}, \mathbf{x}\rangle \mathbf{x} - \Big(\sum_{n=1}^{\infty} \mathbf{x}_n/n^3\Big)(n^3\mathbf{p}_n)_{n\in\mathbb{N}}, \mathbf{w}\Big).
\end{aligned}$$

We conclude that if a spray existed, its quadratic form needs to be given by

$$Q(\mathbf{p}, \mathbf{x}) = 2\langle \mathbf{p}, \mathbf{x}\rangle \mathbf{x} - \Big(\sum_{n=1}^{\infty} \mathbf{x}_n/n^3\Big)(n^3\mathbf{p}_n)_{n\in\mathbb{N}}$$

and this expression is ill defined if $(n^3\mathbf{p}_n)_{n\in\mathbb{N}}$ is not contained in ℓ^2. Hence g does not admit a metric spray.

Exercises

4.3.1 Let $U \subseteq E$, where E is a locally convex space, $0 \in U$ and $f\colon U \to E$ is smooth. Prove that if f is *locally homogeneous of order* p,

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

that is, $f(tx) = t^p f(x)$ for all $(t,x) \in \mathbb{R} \times U$ such that $tx \in U$, then $f(x) = \frac{1}{p!} d^p f(0; x, \ldots, x)$.

4.3.2 Show that the local formula (4.5) defines a spray on a chart domain U. Then show that for any other chart V, the change of charts relates the local formulae obtained in this way to each other, that is, the sprays can be combined to define a spray on TM.

4.3.3 Let S be a spray on a manifold M, (U,φ) and (V,ψ) are two charts of M with change of charts $\tau := \psi \circ \varphi^{-1}$. Prove the following change of charts formulae for the local expression of the spray and the associated bilinear form (see 4.17):

$$S_{V,2}(\tau(x), d\tau(x;v)) = d^2\tau(x;v,v) + d\tau(x; S_{U,2}(x,v)), \tag{4.7}$$

$$B_V(\tau(x); d\tau(x;v), d\tau(x;w)) = d^2\tau(x;v,w) + d\tau(x; B_U(x;v,w)). \tag{4.8}$$

4.3.4 Define the subspace $S := \{\mathbf{x} \in \ell^2 \mid (n^3\mathbf{x}_n)_{n\in\mathbb{N}} \in \ell^2\}$ of ℓ^2. Endow S with the restriction of the weak Riemannian metric $g_{\mathbf{p}}(\mathbf{x},\mathbf{y}) = e^{-\|\mathbf{p}\|^2} \sum_{n=1}^{\infty} \frac{\mathbf{x}_n \mathbf{y}_n}{n^3}$ from Example 4.12. Show that the resulting weak Riemannian metric:

(a) has vanishing geodesic distance;

(b) admits a metric spray and deduce that the (non-)existence of a (smooth) metric spray does not imply the non-degeneracy of the geodesic distance.
Remark: Again I am indebted to C. Maor for pointing this example out to me.

Covariant Derivatives

On a locally convex space E, we can identify vector fields with functions, whence for two vector fields $X,Y \in C^\infty(E,E)$, we can differentiate Y in the direction of X via $X.Y(v) = dY(v, X(v))$ (see Appendix D). This yields again a smooth function from E to E. On a manifold M, the corresponding construction for vector fields would be $TY \circ X \colon M \to T^2M$ which is obviously *not* a vector field. Hence to define a derivative taking two vector fields to vector fields, an additional structure is needed.

4.23 Definition A *covariant derivative* $\nabla \colon \mathcal{V}(M) \times \mathcal{V}(M) \to \mathcal{V}(M)$, $(X,Y) \mapsto \nabla_X Y$ is a bilinear map satisfying the properties

(a) for $f \in C^\infty(M,\mathbb{R}), X, Y \in \mathcal{V}(M)$ we have

1. $\nabla_{fX} Y = f \nabla_X Y$,

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

2. $\nabla_X(fY) = \mathcal{L}_X(f)Y + f\nabla_X Y$ (where $\mathcal{L}_X(f)$ is the Lie derivative (D.3)).

(b) $\nabla_X Y - \nabla_Y X = [X,Y]$ for all $X,Y \in \mathcal{V}(M)$.

We shall show now that every spray induces a covariant derivative. Recall from Appendix D that for a chart (φ,U) we denote by X_φ the local representative of a vector field $X \in \mathcal{V}(M)$.

4.24 Proposition *Given a spray S on M, there exists a unique covariant derivative ∇ such that in a chart (φ,U), the derivative is given by the local formula*

$$(\nabla_X Y)_\varphi(x) = X_\varphi.Y_\varphi(x) - B_U(x; X_\varphi(x), Y_\varphi(x)). \tag{4.9}$$

We call ∇ the covariant derivative associated to the spray *S. Suppressing the indices, the above formula reads $\nabla_X Y = X.Y - B(X,Y)$.*

Proof Define $\nabla_X Y$ locally over U via the formula (4.9). In Exercise 4.3.6 we shall see that this formula has all properties of a covariant derivative for vector fields on U (and all of the defining properties of a covariant derivative localise on U!). Obviously we can repeat the construction for every chart in an atlas $\mathcal{A}$. It suffices now to prove that the family $((\nabla_X Y)_\varphi)_{\varphi\in\mathcal{A}}$ induces a vector field, that is, in view of Lemma D.7, it suffices to prove that the local representatives are related by the change of charts. We will check this for $\tau := \psi\circ\varphi^{-1}$, that is, we prove $d\tau\circ(\nabla_X Y)_\varphi = (\nabla_X Y)_\psi\circ\tau$. Note first that by construction $Y_\psi\circ\tau = d\tau\circ Y_\varphi$ and thus (1.7) yields

$$d(Y_\psi\circ\tau)(x;v) = d(d\tau(x;Y_\varphi)(x;v) = d^2\tau(x;Y_\varphi(x),v) + d\tau(x,dY_\varphi(x;vV)). \tag{4.10}$$

We now apply the change of charts formulae for the spray and associated bilinear form from Exercise 4.3.3:

$$\begin{aligned}
(\nabla_X Y)_\psi(\tau(z)) &= dY_\psi(\tau(z); d\tau(z; X_\varphi(z)) \\
&\quad - B_V(\tau(z); d\tau(z; X_\varphi(z)), d\tau(z; Y_\varphi(z))) \\
&\overset{(4.8)}{=} d(Y_\psi\circ\tau)(z; X_\varphi(z)) - d^2\tau(z; Y_\varphi(z), X_\varphi(z)) - d\tau(z; B_U(z; X_\varphi(z), Y_\varphi(z))) \\
&\overset{(4.10)}{=} d^2\tau(z; Y_\varphi(z), X_\varphi(z)) + d\tau(x, dY_\varphi(x; X_\varphi(z))) \\
&\quad - d^2\tau(z; Y_\varphi(z), X_\varphi(z)) - d\tau(z; B_U(z; X_\varphi(z), Y_\varphi(z))) \\
&= d\tau(z; (\nabla_X Y)_\varphi(z)),
\end{aligned}$$

where we have exploited that the second derivative $d^2\tau(z;\cdot)$ is symmetric by Schwarz' theorem, Exercise 1.3.3. □

4.25 Remark For later use, we observe that the covariant derivative depends only on the values of the field in whose direction we derivate: Let X,Y be

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

vector fields and ∇ be a covariant derivative associated to a spray S. Then the formula (4.9) shows that $\nabla_X Y(p)$ depends only on $X(p)$ but not on the vector field X in a neighbourhood of p.

As soon as we have a covariant derivative associated to a spray, we can define an associated curvature tensor. We recall the definition here for later use.

4.26 Definition Let M be a manifold with a covariant derivative ∇. For vector fields $X, Y, Z \in \mathcal{V}(M)$, we define the linear map

$$R(X,Y)\colon \mathcal{V}(M) \to \mathcal{V}(M) \text{ given by the formula}$$
$$R(X,Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z.$$

Then one can show that R is a trilinear map in the variables X, Y, Z called the *curvature* associated to the covariant derivative; see Exercise 4.3.10. If S is a spray inducing ∇ one also says that R is the curvature of S. Similarly if ∇ is the metric derivative of a (weak) Riemannian metric g, we say that R is the curvature of the metric g.

Curvature (associated to the metric spray) is a fundamental invariant of a Riemannian manifold (Klingenberg, 1995; Lang, 1999; Gallot et al., 2004). Here we mention only that the curvature of certain infinite-dimensional (weak and strong) Riemannian manifolds plays a crucial role in important applications of infinite-dimensional geometry such as Arnold's result (Arnold, 1966) on the practical impossibility of long-term weather forecasts. Also there is an interesting divide between the curvature of strong and weak Riemannian metrics: If (M,g) is a strong Riemannian manifold, the curvature is always (locally) bounded (this follows from the fact that it can be represented as a smooth section into a suitable tensor bundle). However, there are examples of weak Riemannian manifolds with covariant derivative such that the curvature is unbounded (with respect to the norm induced by the metric) locally as well as a multilinear operator on the tangent space at a single point (see Exercise 4.3.11).

4.27 Similar to the construction of a covariant derivative, every spray induces a bundle morphism $K\colon T^2M \to TM$ which is locally (on a chart domain (U,φ)) given by

$$K(x,y,v,w) := (x, w - B_U(x;y,v)). \tag{4.11}$$

This bundle morphism is a linear connection[8] called the *connector* of the spray. By definition of the connector, the associated covariant derivative is $\nabla_X Y = K \circ TY \circ X$.

[8] A linear connection for a bundle $p\colon E \to M$ is given by a bundle map $k\colon TE \to E$ over the bundle projection which is given in bundle trivialisations by bilinear maps as in (4.11). See, for example, Klingenberg (1995, 1.5).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

4.28 Remark If the manifold admits smooth cutoff functions (e.g. if M is finite dimensional; see also §A.4) one can show (Lang, 1999, §VII.2 Theorem 2.4) that every covariant derivative is associated to a smooth spray as in Proposition 4.24.

4.29 Definition A covariant derivative ∇ on a Riemannian manifold (M,g) is called *metric derivative* if for all $X,Y,Z \in \mathcal{V}(M)$, the following equation holds

$$X.g(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z).$$

Here $X.f = df \circ X$ is the derivative of $f\colon M \to \mathbb{R}$ in the direction of the field X.

4.30 Theorem (Lang, 1999, §VIII.4 Theorem 4.1 and 4.2) *Let (M,g) be a strong Riemannian manifold. Then g admits a metric derivative and there exists a unique spray $S\colon TM \to T^2M$, the* metric spray*, whose associated covariant derivative is the metric derivative.*

4.31 Remark To every (strong) Riemannian metric, there is an associated metric spray and a metric derivative. The metric spray turns out to be the one we computed in 4.17 (whence it coincides with the metric spray in Definition 4.21). For a weak Riemannian metric we have no guarantee that a metric spray exists as Example 4.22 shows, but if a metric spray exists, its associated covariant derivative is the metric derivative.

We investigate some examples of Riemannian manifolds for which these objects can be computed explicitly.

4.32 Example (The trivial example) Let $(H,\langle\cdot,\cdot\rangle)$ be a Hilbert space considered as a strong Riemannian manifold. Due to Exercise 4.3.5, the covariant derivative is $\nabla_X Y = X.Y$. This implies that the bilinear form B of the associated spray (aka the Christoffel symbols) needs to vanish, and we see that the connector and the metric spray associated to the metric are

$$\begin{aligned} K\colon T^2H \cong H^4 \to TH \cong H^2, &\qquad (x,u,v,w) \mapsto (x,w),\\ S\colon TH \cong H^2 \to T^2H \cong H^4, &\qquad (x,u) \mapsto (x,u,u,0). \end{aligned}$$

Finally, from the formula for the covariant derivative one sees that the curvature R identically vanishes on H (one also says that H is flat).

4.33 Example (The submanifold example) Recall from Example 4.7 that the Hilbert sphere S_H of a Hilbert space $(H,\langle\cdot,\cdot\rangle)$ is a submanifold of H. This structure turns S_H into strong Riemannian manifold with respect to the pullback metric

$$g_x(V,W) := \langle V,W\rangle \text{ for } W \in T_x S_H = \{y \in H \mid \langle x,y\rangle = 0\}.$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

We shall now describe the covariant derivative of this Riemannian metric. Define the smooth map $\mathrm{pr}\colon S_H \times H \to H$, $(p,v) \mapsto v - \langle p,v\rangle p$, and note that for fixed $p \in S_H$, $\mathrm{pr}_p := \mathrm{pr}(p,\cdot)$ is just the orthogonal projector onto the tangent space T_pS_H. Since the tangent space T_pS_H has been identified with the orthogonal complement of p in H, we may identify vector fields on S_H with smooth maps $X\colon S_H \to H$ which satisfy $\langle X(p),p\rangle = 0$, for all $p \in S_H$. Using these identifications, we can then define a map

$$\nabla^{S_H} : \mathcal{V}(S_H)^2 \to \mathcal{V}(S_H), \nabla^{S_H}_X Y(p) := \mathrm{pr}_p(dY(p,X(p))). \tag{4.12}$$

It is a straightforward computation (Exercise 4.3.7) that ∇^{S_H} defines a covariant derivative on S_H. Let us show now that it is the metric derivative of the pullback metric g. Pick vector fields $X,Y,Z \in \mathcal{V}(S_H)$. We can now compute as follows:

$$\begin{aligned} X.g(Y,Z)(p) &= X.\langle Y,Z\rangle(p) = d\langle Y,Z\rangle(p;X(p)) \\ &= \langle dY(p;X(p));Z(p)\rangle + \langle Y(p),dZ(p;X(p))\rangle \\ &= g(\nabla^{S_H}_X Y,Z)(p) + g(Y,\nabla^{S_H}_X Z)(p), \end{aligned}$$

where the last equality follows from the fact that vector fields on S_H take their image in the orthogonal complement of the base point. Thus ∇^{S_H} is the metric derivative of the pullback metric. The connector and the metric spray are now simply the restrictions of the ones from Example 4.32. There is also an associated formula (called the Gauss equations) relating the curvature of the sphere to the one of the flat ambient space (see Klingenberg, 1995, Example 1.11.6).

The formulae for covariant derivative, spray and connector in Example 4.33 might seem ad hoc and indeed there is not much to see due to the simplicity of the data of the ambient space. However, the whole procedure turns out to be a special case of a formula which relates the covariant derivative of an isometrically immersed submanifold to the covariant derivative of the ambient manifold (see Klingenberg, 1995, Theorem 1.10.3). For the general notion, one needs to define the covariant derivative of vector fields along smooth maps. For now, we shall define this concept only for the special case of a covariant derivative for vector fields along curves (but see Definition 5.4).

4.34 For a C^1-curve $c\colon [a,b] \to M$, we say a C^1 curve $\alpha\colon [a,b] \to TM$ *lifts* c if $\pi_M \circ \alpha = c$. Denote by $\mathrm{Lift}(c)$ the *set of all lifts* of c. Note that the pointwise operations turn $\mathrm{Lift}(c)$ into a vector space on which the smooth functions $C^\infty([a,b],\mathbb{R})$ act by pointwise multiplication.

If X is a (smooth) vector field on M, then for every C^1-curve $c\colon [a,b] \to M$, the curve $X \circ c$ is a lift of c. Note, however, that not every lift of a curve needs to

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

arise as the composition of a vector field and a curve (e.g. if the curve intersects itself, there can be lifts taking different values for the different time parameters associated to the intersection).

4.35 Consider a C^1-curve $c\colon [a,b] \to M$ and a chart (U,φ) of M. We define a local representative $c_U := \varphi \circ c|_{c^{-1}(U)}$ of c in the chart φ. For curves with $\pi_M \circ \alpha = c$ of M, we also define the *principal part* with respect to the chart (U,φ). Namely, we set $T\varphi \circ \alpha|_{\alpha^{-1}(TU)} = (\varphi \circ c|_{\alpha^{-1}(TU)}, \alpha_U) = (c_U(t), \alpha_U(t))$ for some C^1-map α_U.

Furthermore, define for c a curve $\dot{c}\colon [a,b] \to TM$ with the property $\pi_{TM} \circ \dot{c} = c$ as follows: In any chart (U,φ) the principal part of $\dot{c}$ is $(\dot{c})_U(t) := (\varphi \circ c)'(t) = (c_U)'(t)$. Obviously the definition does not depend on the choice of charts and we note that if c is a C^2-curve, then $\dot{c} \in \mathrm{Lift}(c)$. We will later use the same notation for mappings from the circle $\mathbb{S}^1$ with values in a manifold. Note that (up to a harmless identification) the new definition will coincide with the one here; see §5.1.

4.36 Proposition (Lang, 1999, §VIII.3 Theorem 3.1) *Let S be a spray on M with associate bilinear form B and $c \in C^2([a,b],M)$. Then there exists a unique linear map*

$$\nabla_{\dot{c}}\colon \mathrm{Lift}(c) \to \{\gamma \in C([a,b],TM) \mid \pi_M \circ \gamma = c\}$$

which in a chart (U,φ) is given by the formula

$$(\nabla_{\dot{c}}\alpha)_U(t) = \alpha_U'(t) - B_U(c_U(t); c_U'(t), \alpha_U(t)). \tag{4.13}$$

Furthermore, $\nabla_{\dot{c}}$ acts as a derivation on multiplication with C^1-functions φ, that is,

$$\nabla_{\dot{c}}(\varphi\alpha) = \varphi'(t)\nabla_{\dot{c}}\alpha(t) + \varphi(t)\nabla_{\dot{c}(t)}\alpha(t).$$

Proof The proof is similar to Proposition 4.24 and left as Exercise 4.3.9. □

4.37 For a C^2-curve c, a lift γ is *c-parallel* if $\nabla_{\dot{c}}\alpha = 0$. In a chart (U,φ) this is equivalent to $\alpha_U'(t) = B_U(c_U(t); c_U'(t), \alpha_U(t))$. Thanks to Definition 4.19, c is a geodesic for the spray S if and only if $\dot{c}$ is c-parallel, that is, if and only if it satisfies the *geodesic equation*

$$\nabla_{\dot{c}}\dot{c} = 0.$$

4.38 Example From Exercise 4.2.5 we see that a lift $\alpha\colon [a,b] \to H \times H$ of a curve c in a Hilbert space H is c-parallel if and only if its principal part $\mathrm{pr}_2 \circ \alpha$ is constant, that is, the lift corresponds to a curve which is parallel to c in the usual sense of the word.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

From the informal discussion at the end of the last section, we know that a geodesic for a Riemannian manifold should (at least locally) be the shortest path between points which are not far apart. From the presentation in this section this is not apparent. A full proof requires more techniques,[9] whence we just state the following.

4.39 Theorem (Klingenberg, 1995, Theorem 1.9.3) *Let (M,g) be a strong Riemannian manifold, $p \in M$. Then there is an open p-neighbourhood U_p and a constant $\eta > 0$ such that every geodesic starting in U_p of length $< \eta$ is a curve of minimal length between its endpoints.*

Thus a geodesic is always at least locally a curve of minimal length among all curves connecting two (close enough) points on the geodesic.

Exercises

4.3.5 Show that for a Hilbert space $(H,\langle\cdot,\cdot\rangle)$ considered as a strong Riemannian manifold, the metric derivative is given by the usual derivative, that is, $\nabla_X Y = X.Y = dY \circ (\mathrm{id}, X)$. Deduce that geodesics in this case are of the form $at + b$, $a,b \in H$ (see Exercise 4.2.5).
Hint: The metric derivative is unique. If it is $X.Y$, then the bilinear map B of the associated metric spray needs to vanish.

4.3.6 Let (U,φ) be a chart for the manifold M. Show that the local formula (4.9) induces a covariant derivative on $\mathcal{V}(U)$.
Hint: Review Appendix D to prove the statement on the Lie bracket.

4.3.7 Show that (4.12) defines a covariant derivative on S_H.

4.3.8 Consider a Hilbert space $(H,\langle\cdot,\cdot\rangle)$ in the canonical way as a strong Riemannian manifold. Show that for a curve $c \in C^2([a,b],H)$ the covariant derivative $\nabla_{\dot c} f = \dot f_H$ for all $f = (c, f_H) \in \mathrm{Lift}(c)$. Deduce that geodesics in $(H,\langle\cdot,\cdot\rangle)$ are lines in H.

4.3.9 Use Exercise 4.3.3 to work out a proof for Proposition 4.36.

4.3.10 Let M be a manifold with spray S and associated covariant derivative ∇. Work locally in a chart (U,φ) of M, where we write $X_\varphi, Y_\varphi, Z_\varphi$ for the local representatives of vector fields and B_U for the bilinear form associated to the spray.

(a) Establish the following local formula for the curvature:

$$\begin{aligned}(R(X,Y)Z)_\varphi = {} & B_U(B_U(Y_\varphi,Z_\varphi),X_\varphi)) - B_U(B_U(X_\varphi,Z_\varphi),Y_\varphi)) \\ & + d_1B_U(X_\varphi,Z_\varphi;Y_\varphi) - d_1B_U(Y_\varphi,Z_\varphi;X_\varphi).\end{aligned}$$

[9] In Proposition 7.2 we shall see that a curve is geodesic if and only if it extremises energy. As energy bounds length, a geodesics extremises the length (i.e. it is locally of minimal length).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Hint: Recall that B_U takes three arguments $B_U(x; X_\varphi(x), Y_\varphi(x))$ and use Proposition 4.24.

(b) Deduce that $R\colon \mathcal{V}(M)^3 \to \mathcal{V}(M)$ is a trilinear map which satisfies

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0 \text{ (Bianchi identity).}$$

4.3.11 Consider again the subspace $S \subseteq \ell^2$ with the weak Riemannian metric from Exercise 4.3.4. Derive an explicit formula for the curvature $R(X,Y)Z$ and show that the curvature at the point $s = 0$ is unbounded. *Hint:* Everything is local. Since $S \subseteq \ell^2$ is a subspace of a Hilbert space, S admits smooth bump functions. Thus every vector can be continued to a smooth vector field and we write the curvature now for vectors (understanding that they can be continued to vector fields). Let e_i be the vector in ℓ^2 whose only non-zero entry is 1 in the ith place. Set $u_i = \frac{1}{\sqrt{i^3}} e_i$ and work out a formula for $g_0(R(u_i,u_j)u_j|_{s=0}, u_i)$ by exploiting that the e_i are orthonormal with respect to g_0. What happens for $i, j \to \infty$?

Weak Riemannian Metrics with and without Metric Derivative

For a weak Riemannian metric one cannot expect, in general, that there exists a metric derivative associated to the metric. An example of a weak Riemannian metric without an associated covariant derivative can be found in Bauer et al. (2014, p. 12): For a Sobolev-type right invariant metric on a certain subgroup of the diffeomorphism group $\mathrm{Diff}(\mathbb{R})$, the geodesic equation and the covariant derivative do not exist on the subgroup.[10]

4.40 Remark To remedy the problem that the categorisation of weak and strong Riemannian metrics is not sharp enough to capture the existence of covariant derivatives and related important structures for weak Riemannian metrics (see Chapter 5 for an example of a weak Riemannian metric which admits a covariant derivative, connector and metric spray), several authors have proposed a finer classification of infinite-dimensional Riemannian metrics.

We mention here the following concepts and refer interested readers to the original sources for more information.

- Micheli et al. (2013) define a *robust Riemannian metric* as a weak Riemannian metric g on M such that

[10] We will not discuss any details here as this would require at least a discussion of manifold structures on function spaces with non-compact domain. It is worth remarking, though, that the geodesic equation of this metric is related to the non-periodic Hunter–Saxton equation; see Bauer et al. (2014, p. 12 Theorem).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(a) the associated metric derivative exists, and
(b) the Hilbert space completions $\overline{T_xM}^{g_x}$ form a smooth vector bundle $\bigcup_{x\in M}\overline{T_xM}^{g_x}$ whose trivialisations extend the trivialisations of the bundle TM.

In particular, this setting is strong enough to enable certain calculations of curvature for the weak Riemannian manifold.

- Stacey (2008) strengthens the notion of a weak Riemannian structure via an additional structure: As observed in 4.4, one of the main differences between the strong and the weak setting is the failure of the mapping $\flat\colon TM \to T^*M$ to be an isomorphism. While this cannot be directly remedied, requiring a so-called *co-orthogonal structure* allows one to obtain a map replacing the inverse $\sharp$ (which does not exist) of $\flat$ by a mapping with dense image. Exploiting the additional structure it is possible to define a Dirac operator on loop spaces.

We refer to §5.1 for computations of metric derivatives for the (weak) L^2- and H^1-metrics.

4.4 Geodesic Completeness and the Hopf–Rinow Theorem

In this section, we investigate geodesic completeness of infinite-dimensional Riemannian manifolds and the Hopf–Rinow theorem. We will see that this theorem fails in infinite-dimensional geometry as it is built on top of (local) compactness of the underlying manifold. Let us first recall some definitions.

4.41 Definition Let (M,g) be a (weak) Riemannian manifold. Then M

(a) is *metrically complete*, if the geodesic distance turns M into a complete metric space (i.e. Cauchy sequences with respect to the geodesic distance converge);
(b) is *geodesically complete* if every geodesic can be continued for all time;
(c) *has minimising geodesics* if for every two points in the same connected component of M, there exists as length minimising geodesic c connecting the points (i.e. for a,b there is a geodesic c with $\mathrm{Len}(c) = \mathrm{dist}(a,b)$).

In finite dimensions the Hopf–Rinow theorem holds (see Klingenberg, 1995, Theorem 2.1.3).

4.42 Theorem (Hopf–Rinow) *Let (M,g) be a* finite-dimensional *Riemannian manifold. Then M is metrically complete if and only if it is geodesically complete. Moreover, if M is metrically complete, it has minimising geodesics.*

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

In infinite-dimensional settings Theorem 4.42 does not hold. What remains true, however, is that metrical completeness of a strong Riemannian manifold implies geodesic completeness.

Metrically complete implies geodesically complete Let $c_X\colon J \to M$ be a geodesic with $\dot{c}(0) = X \in T_pM$. We argue by contradiction and assume that $\sup J = t^+ < \infty$. We can reparametrise c_X such that $\operatorname{Len}(c_X|_{[r,s]}) = |r-s|$ for all $r, s \in J$. Pick a Cauchy sequence $(t_n)_{n\in\mathbb{N}}$ with $t_n \nearrow t^+$. As

$$\operatorname{Len}(c_X(t_n), c_X(t_m)) \leq |t_n - t_m|,$$

we see that the points $(c(t_n))_{n\in\mathbb{N}}$ form a Cauchy sequence with respect to the geodesic distance, whence they converge towards some limit $q \in M$. Pick now $\varepsilon > 0$ so small that $B_\varepsilon^{\mathrm{dist}}(q)$ is the domain of Riemannian normal coordinates (Lang, 1999, §VIII.6. Theorem 6.4). For $t^+ - t_0 < \varepsilon/2$ there is a geodesic γ with $\dot{\gamma}(0) = \dot{c}(t_0)$. Note that by uniqueness of geodesics $\gamma(t) = c(t+t_0)$. Since γ is contained in the normal coordinates around q, the domain of definition of γ contains at least $[-\varepsilon/2, \varepsilon/2]$, whence c can be extended beyond t^+. □

In infinite dimensions the following example shows that metric and geodesic completeness do not imply existence of minimising geodesics. This is mainly a consequence of the lack of local compactness as the generalised version of the Hopf–Rinow theorem (in the context of metric spaces) shows see Bridson and Haefliger (1999, Proposition 3.7).

4.43 Example (Grossman's ellipsoid (McAlpin, 1965)) Let $H = \ell^2$ be the Hilbert space of square summable sequences with the orthonormal basis $(\mathbf{e}_n)_{n\in\mathbb{N}}$ (where $\mathbf{e}_n$ is the sequence with a 1 in the nth place and 0s everywhere else). Recall that the inner product on ℓ^2 is $\langle (x_n)_n, (y_n)_n\rangle = \sum_n x_n y_n$. We define $a_1 = 1$ and $a_n = 1 + 2^{-n}$ for $n \geq 2$ and consider the ellipsoid

$$E := \left\{ (x_n)_{n\in\mathbb{N}} \in \ell^2 \,\middle|\, \sum_{n\in\mathbb{N}} \frac{x_n^2}{a_n^2} = 1 \right\}.$$

Defining the smooth diffeomorphism $F\colon H \to H$, $(x_n)_{n\in\mathbb{N}} \mapsto (a_n x_n)_{n\in\mathbb{N}}$, we see that the ellipsoid is the image of the Hilbert sphere $E = F(S_H)$. As the Hilbert sphere is a submanifold of H by Example 1.34, so is the ellipsoid by Exercise 4.4.2. We endow the ellipsoid with the strong Riemannian metric induced by the embedding $E \subseteq H$ whence it becomes a strong Riemannian manifold. Note also that S_H is a Riemannian manifold with respect to the induced metric. Thanks to Klingenberg (1995, 1.10.13(iii)) geodesics in this Riemannian manifold are given by great circles and this shows that S_H is complete (we shall study such a great circle in the next step). Note that by

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

definition of F we have $F(r\mathbf{e}_1) = r\mathbf{e}_1$ for all $r \in \mathbb{R}$. Moreover, if γ is a path in S_H connecting $\mathbf{e}_1$ to $-\mathbf{e}_1$, then $F \circ \gamma$ is a path connecting $\mathbf{e}_1$ and $-\mathbf{e}_1$ in E. As F is a diffeomorphism every path connecting $\mathbf{e}_1$ and $-\mathbf{e}_1$ in E arises in this way. Now, thanks to Exercise 4.4.2(d), we can apply Theorem 6.9 of Lang (1999, §VIII.6) which also implies that E is complete (and geodesically complete).

We shall now prove that $\operatorname{dist}_E(\mathbf{e}_1, -\mathbf{e}_1) = \pi$, but there exists no path realising the minimal distance. Moreover, we shall prove in Exercise 4.4.2 that the half circle $\gamma(t) = \cos(\pi t)\mathbf{e}_1 + \sin(\pi t)\mathbf{e}_2, t \in [0,1]$ is a geodesic connecting $\mathbf{e}_1$ and $-\mathbf{e}_1$, whence $\operatorname{dist}(\mathbf{e}_1, -\mathbf{e}_1)$ in S_H is π. Consider now an arbitrary $\gamma\colon [0,1] \to H, \gamma(t) = (\gamma_n(t))_{n\in\mathbb{N}}$ in S_H connecting $\mathbf{e}_1$ and $-\mathbf{e}_1$. For the length of the paths in S_H and E we obtain as a special case of Exercise 4.4.3 the following:

$$\pi \le \operatorname{Len}(\gamma) = \int_0^1 \sqrt{\sum_{n\in\mathbb{N}} \dot\gamma_n(t)^2}\,\mathrm{d}t \le \int_0^1 \sqrt{\sum_{n\in\mathbb{N}} a_n^2 \dot\gamma_n(t)^2} = \operatorname{Len}(F(\gamma)). \tag{4.14}$$

By definition of F, we have equality $\operatorname{Len}(\gamma) = \operatorname{Len}(F\gamma)$ if and only if $\dot\gamma_n(t) = 0$ for $n \ge 2$. However, the only curve starting in $\mathbf{e}_1$ satisfying this is the constant curve. Hence we have for all curves joining $\mathbf{e}_1$ and $-\mathbf{e}_1$ the strict inequality $\operatorname{Len}(F \circ \gamma) > \pi$. Considering now the half ellipse $\gamma_n(t) = F(\cos(\pi t)\mathbf{e}_1 + \sin(\pi t)\mathbf{e}_n), t \in [0,1]$ joining $\mathbf{e}_1$ and $-\mathbf{e}_1$ in the $(\mathbf{e}_1, \mathbf{e}_n)$-plane, then

$$\begin{aligned}\operatorname{Len}(\gamma_n) &= \int_0^1 \pi \sqrt{\sin^2(\pi t) + (1+2^{-n})^2 \cos^2(\pi t)}\,\mathrm{d}t \le \sqrt{1+2^{-n}}\,\pi \to \pi \\ &= \operatorname{dist}_E(\mathbf{e}_1, -\mathbf{e}_1).\end{aligned}$$

This illustrates the failure of the Hopf–Rinow theorem in infinite dimensions. Note that explicit counterexamples are known for the other relations (between completeness, geodesic completeness and existence of minimising geodesics) established in the Hopf–Rinow theorem. See Atkin (1975, 1997) for more information.

4.44 Remark It should be mentioned that parts of the Hopf–Rinow theorem can also be salvaged in the infinite-dimensional setting. However, then certain assumptions on the curvature of the Riemannian manifold are needed. For a manifold with seminegative curvature, geodesic completeness implies completeness; see for example, Lang (1999, §IX.3).

Exercises

4.4.1 Let c be a geodesic in a Riemannian manifold (M, g) with covariant derivative ∇.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(a) Show that if $\varphi\colon \mathbb{R} \to \mathbb{R}$, $t \mapsto at + b$, for $a, b \in \mathbb{R}$, then the reparametrisation $c \circ \varphi$ is also a geodesic.

(b) Let $\mathrm{Len}(c|_{[a,b]}) = L|a - b|$. Show that there is a reparametrisation ψ such that $\mathrm{Len}(c \circ \psi|_{[c,d]}) = |c - d|$ for all c, d in the domain of the reparametrised geodesic.

4.4.2 Check various details for Example 4.43.

(a) Show that the mapping $F\colon H \to H$, $F(\sum x_n e_n) = \sum a_n x_n e_n$ is a smooth map with smooth inverse.
Hint: Observe that F is linear and use the idea that limits and series can be exchanged if the series is dominated by a convergent majorant.

(b) Let $S \subseteq M$ be a split submanifold and $F\colon M \to N$ a diffeomorphism. Show that $F(S)$ is a split submanifold of N.
Hint: See Lemma 1.61.

(c) The covariant derivative on S_H as a submanifold of H with the induced metric is given as $\nabla = \mathrm{pr} \circ \nabla^H$, where ∇^H is the covariant derivative on H and pr the projector onto the tangent space of S_H (see Example 4.33). Prove that the half-circle $\gamma(t) = \cos(\pi t)e_1 + \sin(\pi t)e_2$ is a geodesic in S_H.

(d) Show that for all $x, v \in H$ we have $\|T_x F(v)\| \geq \|v\|$.

4.4.3 Generalise (4.14) in the following way: If $f\colon (M, g) \to (N, h)$ is a map between Riemannian manifolds such that there exists a constant $C > 0$ with

$$\sqrt{h_{f(\pi(v))}(Tf(v), Tf(v))} \geq C \sqrt{g_{\pi(v)}(v, v)} \qquad \text{for all } v \in TM,$$

show that for a piecewise C^1-path $\gamma\colon [a, b] \to M$ we have $\mathrm{Len}(f \circ \gamma) \geq C\mathrm{Len}(\gamma)$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

5

Weak Riemannian Metrics with Applications in Shape Analysis

In this chapter, we study in detail the (weak) L^2-metric on spaces of smooth mappings. Its importance stems from the fact that this metric and its siblings, the Sobolev H^s-metrics, are prevalent in shape analysis. It will be essential for us that geodesics with respect to the L^2-metric can be explicitly computed. Before we look into the specifics, let us clarify what we mean here by shape and shape analysis. Shape analysis seeks to classify, compare and analyse shapes. As a mathematical discipline, shape analysis goes back to the classical works by D'Arcy Thompson (1942) (originally published in 1917). In recent years there has been an explosion of applications in shape analysis to diverse areas such as computer vision (Celledoni et al., 2016), medical imaging, registration of radar images and many more (see Bauer et al., 2014, for an exposition). There are different mathematical settings as to what is meant by a shape and what kind of data describes it. Popular choices are, for example,

- points;
- curves (or surfaces) in Euclidean space of manifolds;
- level sets of functions; and
- images.

Another typical feature in (geometric) shape analysis is that one wants to remove superfluous information from the data. For example, in the comparison of shapes, rotations, translations, scalings and reflections are typically disregarded as being inessential differences. Conveniently, these inessential differences can mostly be described by actions of suitable Lie groups (such as the rotation and the diffeomorphism groups). This hints at the general process of constructing an (infinite-dimensional) manifold of shapes: One starts out with an (infinite-dimensional) manifold of data (e.g. smooth curves) called the preshape space. Then the undesirable information is removed by quotienting out

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

suitable group actions (e.g. if reparametrisation invariance of the shapes is desired, quotient out a suitable diffeomorphism group). The resulting quotient is then called shape space and one seeks to construct suitable tools (such as a Riemannian metric) to compare, classify and analyse the objects in shape space.

In the present chapter we will restrict our attention to shapes which arise as images of smooth curves which take their values in $\mathbb{R}^2$. The pre-shape space will thus be the infinite-dimensional manifold of smooth immersions (from the circle) with values in the two-dimensional space. As the objective is to compare images of these curves, we need to remove the specific parametrisation of the immersion. Hence we pass to shape space by quotienting out an action of the diffeomorphism group on the immersions (modelling the reparametrisation). Our aim is then to construct a suitable Riemannian metric on shape space which will allow us to compare shapes using the geodesic distance induced by it.

5.1 The L^2-metric and Its Cousins

Having now discussed ideas to remedy the problems in the (in general) ill-behaved weak Riemannian setting, we now consider several weak Riemannian metrics which admit metric derivatives, spray and so on. The metrics which we will consider are on one hand the L^2-metric. We will see that the L^2-metric admits a spray, a connector and a covariant derivative albeit being only a weak Riemannian metric. This metric will play a decisive role in the investigation of shape analysis in this chapter. The construction for the L^2-metric follows the argument first presented in Bruveris (2018). Finally, we will briefly describe a Sobolev-type metric whose covariant derivative can explicitly be given and is of independent interest. Before we begin, let us set some conventions concerning the integration of functions on manifolds.

On Integration of Functions on $\mathbb{S}^1$ The unit circle $\mathbb{S}^1$ can be parametrised (up to the double endpoint) by $\theta\colon [0,2\pi] \to \mathbb{S}^1$, $t \mapsto (\cos(t), \sin(t))$. We now abuse notation and denote by θ both the parameter and the parametrisation. If you have not seen integration theory on submanifolds of $\mathbb{R}^d$ (e.g. Lang, 1999, XVI), this implies that for a continuous $f\colon \mathbb{S}^1 \to \mathbb{R}^d$, the integral on $\mathbb{S}^1$ satisfies

$$\int_{\mathbb{S}^1} f(\theta)\mathrm{d}\theta = \int_0^{2\pi} f(\cos(t), \sin(t))\mathrm{d}t,$$

whence it can be computed as a usual one-dimensional integral. Further, for differentiable maps $c\colon \mathbb{S}^1 \to \mathbb{R}$ we write $c'(\theta) \coloneqq T_\theta c(1)$, where $1 \in T_\theta\mathbb{S}^1 \cong \mathbb{R}$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(the notation was previously reserved for curves and we justify it as $T_{\theta(t)}c(1)$ is (up to the linear isomorphism $T_t\theta$) given by the curve differential $(c\circ\theta)'$).

5.1 Remark None of the techniques employed in this chapter depend on the compact manifold $\mathbb{S}^1$, but we can thus skip a discussion of integration against a volume form on a compact manifold. However, we remark that all of the results in this section carry over if we replace $\mathbb{S}^1$ by an arbitrary compact manifold (we invite the reader to check this for themselves).

As in Example 4.3 consider the space $C^\infty(\mathbb{S}^1,M)$ for a strong Riemannian manifold (M,g) with the L^2-metric

$$g^{L^2}_c(f,g)=\int_{\mathbb{S}^1} g_{c(\theta)}(f(\theta),g(\theta))\mathrm{d}\theta,\qquad f,g\in T_cC^\infty(\mathbb{S}^1,M). \tag{5.1}$$

We shall show that this weak Riemannian metric admits a metric spray and a metric derivative. It will turn out that all the relevant objects can be lifted from the target manifold to the manifold of mappings.

5.2 Let (M,g) be a strong Riemannian manifold with metric spray S, connector K and metric derivative ∇. Since (M,g) is a strong Riemannian manifold, it admits a local addition (Michor, 1980, Lemma 10.2), whence the manifold structure on $C^\infty(\mathbb{S}^1,M)$ is canonical and, in addition, $TC^\infty(\mathbb{S}^1,M)\cong C^\infty(\mathbb{S}^1,TM)$ and $T^2C^\infty(\mathbb{S}^1,M)\cong C^\infty(\mathbb{S}^1,T^2M)$ (see Appendix C). Then the pushforward of the spray and the connector

$$\begin{aligned} K_*\colon C^\infty(\mathbb{S}^1,T^2M)&\to C^\infty(\mathbb{S}^1,TM),\quad q\mapsto K\circ q,\\ S_*\colon C^\infty(\mathbb{S}^1,TM)&\to C^\infty(\mathbb{S}^1,T^2M),\quad v\mapsto S\circ v \end{aligned}$$

are smooth mappings by Corollary 2.19. Moreover, the identification of the tangent bundles immediately shows that S_* is a spray on $C^\infty(\mathbb{S}^1,M)$ and that K_* is a connector.

5.3 Proposition *In the situation of* 5.2*, the connector associated to* S_* *is* K_*.

Proof By definition, the connector of S_* is uniquely determined by the associated map B_{S_*}. If we can show that B_{S_*} coincides with the pushforward B_* of the map B associated to S, then the definition of the connectors yields that the connector of S_* is K_*. To compute B_{S_*} we need to compute $d^2_2(S_*)_{O,2}((x,0);\cdot)$, where $(S_*)_{O,2}$ is a local reprensentative of S_* (see 4.17). Since the bundle trivialisations are complicated for manifolds of mappings, we apply instead the exponential law to see that we can just take partial derivatives of

$$(S_*)^\wedge\colon C^\infty(\mathbb{S}^1,TM)\times\mathbb{S}^1\to T^2M,\quad (h,\theta)\mapsto S(h(\theta)).$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Observe that we only need to take partial derivatives with respect to the first component; thus $\mathbb{S}^1$ is simply a parameter set and the formula holds since it can be checked pointwise (see Exercise 5.1.1). □

We use the connector K of the covariant derivative on (M,g) to define a covariant derivative for vector fields along smooth maps.

5.4 Definition Let N be a smooth manifold and $F\colon N \to M$ be a smooth map to a Riemannian manifold (M,g). A mapping $s \in C^\infty_F(N,TM) = \{s \in C^\infty(N,TM) \mid \pi_M \circ s = F\}$ is called a *vector field along* F. Assume that (M,g) admits a metric spray with associated connector K. For $X \in \mathcal{V}(N)$, we define

$$\nabla^g_X s := K \circ Ts \circ X\colon N \to TM. \tag{5.2}$$

Note that the construction in Definition 5.4 is a generalised version of Proposition 4.36. If ∇ is the metric derivative of a Riemannian manifold which admits a metric spray then the *covariant derivative along* $F\colon N \to M$ also satisfies a version of 4.29 (see e.g. Klingenberg, 1995, Proposition 1.8.14):

$$\begin{aligned} &X.g(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z),\\ &Y,Z \text{ vector field along } F, X \in \mathcal{V}(N). \end{aligned} \tag{5.3}$$

We will exploit the exponential law for the canonical manifold $C^\infty(\mathbb{S}^1,M)$. To $s \in C^\infty(N,C^\infty(\mathbb{S}^1,TM))$ we associated the map $s^\wedge \in C^\infty(N \times \mathbb{S}^1, M)$.

5.5 Remark We face a problem as in Example 3.25: The exponential law suggests working with vector fields on $N \times \mathbb{S}^1$, while the covariant derivative in (5.2) is only defined for vector fields on N. Luckily, vector fields on product manifolds are products of vector fields on the parts. Hence we extend $X \in \mathcal{V}(N)$ to a vector field on $N \times \mathbb{S}^1$ by the zero vector field on $\mathbb{S}^1$ via $X \times \mathbf{0}_{\mathbb{S}^1} \in \mathcal{V}(N \times \mathbb{S}^1)$. This allows us to obtain vector fields on the correct manifold on which we can now extend the covariant derivative of N.

5.6 (A covariant derivative on $C^\infty(\mathbb{S}^1,M)$) Choosing $N = C^\infty(\mathbb{S}^1,M)$ we define via (5.2) a map $\nabla^g_{X\times\mathbf{0}_{\mathbb{S}^1}} s^\wedge \in C^\infty(C^\infty(\mathbb{S}^1,M)\times\mathbb{S}^1, TM)$ and set

$$\nabla_X s := (\nabla^g_{X\times\mathbf{0}_{\mathbb{S}^1}} s^\wedge)^\vee \in C^\infty(C^\infty(\mathbb{S}^1,M), C^\infty(\mathbb{S}^1,TM)). \tag{5.4}$$

Observe that the identification $TC^\infty(\mathbb{S}^1,M) \cong C^\infty(\mathbb{S}^1,TM)$ allows us to identify $s \in \mathcal{V}(C^\infty(\mathbb{S}^1,M)) \subseteq C^\infty(C^\infty(\mathbb{S}^1,M), C^\infty(\mathbb{S}^1,TM))$. Computing with the help of Exercises 5.1.2 and 2.2.2,

$$\begin{aligned} \pi_{C^\infty(\mathbb{S}^1,M)} \circ \nabla_X s &= (\pi_M)_*(\nabla^g_{X\times\mathbf{0}_{\mathbb{S}^1}} s^\wedge)^\vee = (\pi_M \circ (\nabla^g_{X\times\mathbf{0}_{\mathbb{S}^1}} s^\wedge)^\vee\\ &= (\pi_M \circ K \circ T(s^\wedge) \circ (X \times \mathbf{0}_{\mathbb{S}^1}))^\vee = (\pi_M \circ K \circ (Ts \circ X)^\wedge)^\vee\\ &= (\pi_M \circ K)_*(Ts \circ X) = \mathrm{id}_{C^\infty(\mathbb{S}^1,M)}. \end{aligned}$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Hence (5.4) induces a bilinear map $\nabla\colon \mathcal{V}(C^\infty(\mathbb{S}^1,M))^2 \to \mathcal{V}(C^\infty(\mathbb{S}^1,M))$.

The construction 5.6 certainly looks quite messy; however, we remark that it is indeed a very natural covariant derivative we obtain in this way. Namely, for any smooth curve $c\colon [a,b] \to C^\infty(\mathbb{S}^1,M)$ and smooth lift $\alpha \in \operatorname{Lift}(c)$ we apply the exponential law to obtain smooth maps $c^\wedge\colon [a,b]\times\mathbb{S}^1 \to M$ and $\alpha^\wedge\colon [a,b]\times\mathbb{S}^1 \to TM$. Now the covariant derivative ∇ from 5.6 is related to the covariant derivative ∇^g on M by the following formula:

$$\nabla_{\dot c}\alpha(\cdot)(x) = \nabla^g_{\dot c^\wedge(\cdot,x)}\alpha(\cdot,x) \qquad \text{for all } x\in\mathbb{S}^1. \tag{5.5}$$

We relegate the verification of (5.5) to Exercise 5.1.2. However, the point is that, despite the technical difficulties in defining the covariant derivative, it can be viewed just as a lifting of the covariant derivative of the target manifold M.

5.7 Proposition *For the map $\nabla\colon \mathcal{V}(C^\infty(\mathbb{S}^1,M))^2 \to \mathcal{V}(C^\infty(\mathbb{S}^1,M))$ and $X,Y\in\mathcal{V}(C^\infty(\mathbb{S}^1,M))$, the formula*

$$\nabla_X Y = K_* \circ TY \circ X \tag{5.6}$$

holds. Since K_* is a linear connector, ∇ is a covariant derivative (with associated connector $K_*\colon C^\infty(\mathbb{S}^1,T^2M) \to C^\infty(\mathbb{S}^1,TM)$). In particular, (5.4) is the covariant derivative associated to the spray S_*.

Proof By Proposition 5.3 the connector K_* is the associated connector to the spray S_*. Hence it suffices thus to prove (5.6). We have essentially done this already as the exponential law and Exercise 5.1.2 yield

$$\begin{aligned}(\nabla_X s)^\wedge &= \nabla^g_{X\times\mathbf{0}_{\mathbb{S}^1}} s^\wedge = K\circ Ts^\wedge\circ(X\times\mathbf{0}_{\mathbb{S}^1}) = K\circ(Ts\circ X)^\wedge\\ &= (K_*\circ Ts\circ X)^\wedge.\end{aligned}$$

□

5.8 Proposition *Let $(H,\langle\cdot,\cdot\rangle)$ be a Hilbert space with a strong Riemannian metric g (not necessarily $g=\langle\cdot,\cdot\rangle$). For the L^2-metric (5.1) on $C^\infty(\mathbb{S}^1,H)$, the metric derivative is the covariant derivative ∇ from 5.6.*

Proof Recall that $TH\oplus TH = H\times H\times H$ and we can consider g as a smooth map of three variables. Let's agree that the first component represents the base point in the bundle. Since S is the spray of the metric g, the associated bilinear form B satisfies for all $X,Y,Z\in H$ the relation

$$-2g(p,(B(p;X,Y),Z) = d_1g(p,Y,Z;X) + d_1g(p,Z,X;Y) - d_1g(p,X,Y;Z); \tag{5.7}$$

see Lang (1999, §VIII.4 Theorem 4.2) or Klingenberg (1995, Theorem 1.8.11). Hence if we can compute that an identity such as (5.7) holds for the bilinear

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

form associated to S_* (which is given as pushforward with B) with respect to g^{L^2}, this implies that g^{L^2} admits an associated bilinear form (aka Christoffel symbols). Thus if we know the Christoffel symbols, we can compute the connector giving the metric derivative. In this case, 5.6 shows that the connector is K_*, whence S_* is the metric spray and ∇ the metric derivative of the L^2-metric (see also Theorem 4.30).

Let us now establish the desired analogue of (5.7) for the L^2-metric. The integral operator $\int_{\mathbb{S}^1} : C^\infty(\mathbb{S}^1,\mathbb{R}) \to \mathbb{R}$ is continuous linear. Up to the identification C.13, the derivative of $g^{L^2} = \int_{\mathbb{S}^1} \circ g_*$ in a direction is thus given by the pointwise derivative, that is,

$$d_1 g^{L^2}(c,h,k;\xi) = \int_{\mathbb{S}^1} d_1 g(c(\theta),h(\theta),k(\theta);\xi(\theta))\mathrm{d}\theta.$$

We apply this observation to the right-hand side of (5.7) and recall from Proposition 5.3 that the associated bilinear form B_{S_*} of S_* is given by the pushforward of the associated bilinear form B of S. Together this yields

$$\begin{aligned} &d_1 g^{L^2}(c,Y,Z;X) + d_1 g^{L^2}(c,Z,X;Y) - d_1 g^{L^2}(c,X,Y;Z) \\ &= \int_{\mathbb{S}^1} d_1 g(c,Y,Z;X) + d_1 g(c,Z,X;Y) - d_1 g(c,X,Y;Z) d\theta \\ &= \int_{\mathbb{S}^1} -2g(c(\theta),B(c(\theta);X(\theta),y(\theta),Z(\theta))\mathrm{d}\theta = -2g^{L^2}(c,B_{S_*}(c;X,Y),Z). \quad \square \end{aligned}$$

5.9 Remark (a) Using the point evaluations of $TC^\infty(\mathbb{S}^1,M) \cong C^\infty(\mathbb{S}^1,TM)$, it is possible to directly describe the metric derivative of the L^2-metric by looking at it pointwise evaluated. While this allows one to avoid the construction in 5.6, one is then left with lots of localisation arguments to establish Proposition 5.8. We refer to Maeda et al. (2015, Lemma 2.1) for more information.

(b) All of the above computations for spray, connector and covariant derivative did not exploit that the source manifold was $\mathbb{S}^1$. Thus via the same proof one can obtain a spray, connector and covariant derivative for the L^2-metric on $C^\infty(K,M)$ for any compact manifold K.

(c) Since (5.7) can be formulated in any local chart, a more involved argument works for any strong Riemannian manifold M. However, using the Nash embedding theorem, Proposition 5.8 generalises directly to all finite-dimensional Riemannian manifolds as Bruveris (2018, Theorem 4.1) shows.

We have identified the covariant derivative of the L^2-metric, and thus we can describe geodesics of the L^2-metric: A geodesic between $f,g \in C^\infty(\mathbb{S}^1,H)$ is

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

a path of curves which is pointwise for every $\theta \in \mathbb{S}^1$ a geodesic in H between $f(\theta)$ and $g(\theta)$. Let us explicitly compute this in the case where the inner product of the Hilbert space gives us the Riemannian metric.

5.10 Example Consider $(H, \langle\cdot,\cdot\rangle)$ as a strong Riemannian manifold. Recall that the metric spray and the covariant derivative were computed for this metric in Example 4.32: With the identification $T^k H = H^{2^k}, k \in \mathbb{N}$ we obtain $S(x,v) = (x,v,v,0)$ and $\nabla_X Y = X.Y$ for suitable paths and their lifts $\nabla_{\dot c} h = \dot h$. Now endow $C^\infty(\mathbb{S}^1, H)$ with the L^2-metric and pick $p \in C^\infty(\mathbb{S}^1, H)$. We now compute the geodesics $c\colon J \to C^\infty(\mathbb{S}^1, H)$ starting at p with derivative $\dot c = X$. Now identify the first and second derivatives as $\dot c(t) = (c(t), c'(t)) \in C^\infty(\mathbb{S}^1, H\times H)$ and $\ddot c(t) = (c(t), c'(t), c'(t), c''(t))$. The exponential law allows us to compute the geodesic equation with respect to the spray as

$$(c(t), c'(t), c', c''(t))^\wedge = (\ddot c(t))^\wedge = (S_*(c(t), c'(t)))^\wedge = S\left(c^\wedge(t,\cdot), \frac{\partial}{\partial t} c^\wedge(t,\cdot)\right).$$

Evaluating in $\theta \in \mathbb{S}^1$ one immediately sees that $c''(t) = 0$, for all t. Hence, from the usual rules of calculus we deduce that $c(t)(\theta) = tX(\theta) + p(\theta)$. A geodesic between $f, g\colon \mathbb{S}^1 \to H$ does thus always exist and is described by a map $\gamma\colon [0,1] \times \mathbb{S}^1 \to H$ such that for every $\theta \in \mathbb{S}^1$ we have $\gamma(t,\theta) = (1-t) f(\theta) + t g(\theta)$. In other words, geodesics in this example are given by pointwise linear interpolation between the two functions.

We shall further investigate the geodesic equation of a generalised version of the L^2-metric on the open submanifold $\mathrm{Diff}(M) \subseteq C^\infty(M, M)$ in Chapter 7. We can also use our knowledge of the L^2-metric to derive other interesting examples of metric derivatives for certain weak Riemannian metrics. Note, however, that the following example requires a more in-depth knowledge of Riemannian geometry (e.g. the Hodge Laplacian, Definition E.16 and curvature Definition 4.26).

5.11 Example (Maeda et al., 2015) Let (M, g) be a finite-dimensional Riemannian manifold with metric derivative ∇^g. We recall that every (finite-dimensional) Riemannian manifold has a (Hodge)Laplacian $\Delta = \mathrm{dd}^* + \mathrm{d}^*\mathrm{d}$ associated to the metric (see Lang, 1999, p. 423) and a curvature tensor R (see Definition 4.26).

Now consider the loop space $LM := C^\infty(\mathbb{S}^1, M)$. By Remark 5.9(c) we know that the L^2-metric admits a metric derivative ∇. We will now use the notation of lifts and covariant derivative along curves on $[0, 2\pi]$ in the context of curves on $\mathbb{S}^1$ (implicitely identifying elements in LM with curves on $[0, 2\pi]$ by composing with $(\sin, \cos)\colon [0, 2\pi] \to \mathbb{S}^1$). Hence for a smooth curve $\gamma \in LM$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

we have $T_\gamma C^\infty(\mathbb{S}^1, M) = \mathrm{Lift}(\gamma) \cap C^\infty(\mathbb{S}^1, TM)$. We can then endow every tangent space of $C^\infty(\mathbb{S}^1, M)$ with the H^1-inner product:

$$g_\gamma^{H^1}(X,Y) := \frac{1}{2\pi}\int_0^{2\pi} g_{\gamma(\theta)}\left((1+\Delta)X(\theta), Y(\theta)\right) \mathrm{d}\theta. \tag{5.8}$$

Rewriting the codifferential in the above formula and exploiting duality with respect to the metric g, one can show that the inner product (5.8) describes the sum of the L^2-inner products of the lift of γ and its first derivative. We can thus leverage Exercise 4.1.4(a), where we have seen that the L^2-metric is a (weak) Riemannian metric. Differentiation is continuous linear (on the tangent space of $C^\infty(\mathbb{S}^1, M)$), whence the H^1-metric is a weak Riemannian metric. Now a (non-trivial!) computation shows that the metric derivative of the H^1-metric is intimately connected to the curvature tensor, the metric derivative of g and the metric derivative of the L^2-metric. Namely, Maeda et al. (2015, Theorem 2.2) provide for $X, Y \in T_\gamma C^\infty(\mathbb{S}^1, M)$ the following formula for the metric derivative $\nabla_X^{H^1} Y(\gamma)$:

$$\begin{aligned}\nabla_X Y + \frac{1}{2}(1+\Delta)^{-1}\Big(&-\nabla^g_{\dot c}(R(X,\dot c)Y) - R(X,\dot c)\nabla^g_{\dot c} Y - \nabla^g_{\dot c}(R(Y,\dot c))X\\ &-R(Y,\dot c)\nabla^g_{\dot c} X + R(X,\nabla^g_{\dot c} Y)\dot c - R(\nabla^g_{\dot c} X, Y)\dot c\Big),\end{aligned}$$

where $\nabla^g_{\dot c}(R(X,\dot c)Y)$ denotes the lift of γ whose value at $\theta \in \mathbb{S}^1$ is given by the formula $-\nabla^g_{\dot c(\theta)}(R(X(\theta), \dot c(\theta))Y(\theta))$. While the above formula looks daunting and we do not attempt to unravel its meaning here, we would like to mention that it can be used to connect the Riemannian geometry of the H^1-metric to pseudodifferential operators acting on a trivial bundle over the circle. This link then yields information on Chern–Simons classes on the tangent bundle of the loop space. We refer the interested reader to Maeda et al. (2015) for more information.

Exercises

5.1.1 Let $S\colon TM \to T^2M$ be a spray on M with $K\colon T^2M \to TM$ its associated connector.

(a) Prove that S_* is a spray on $C^\infty(\mathbb{S}^1, M)$.
Hint: Use that $TC^\infty(\mathbb{S}^1, M) \cong C^\infty(\mathbb{S}^1, TM)$ identifies $T(p_*)$ with $(Tp)_*$ for smooth maps. For the quadratic condition review the effect of the diffeomorphism on the fibres.

(b) Check that K_* is a connector.

(c) Show that the second derivative of the vertical part of S_* is the pushforward of the second derivative of the vertical part of S (see Proposition 5.3).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

5.1.2 Prove for $s \in C^\infty(N, C^\infty(\mathbb{S}^1, M))$ and $X \in \mathcal{V}(N)$ the identity

$$Ts \circ X = (Ts^\wedge \circ (X \times \mathbf{0}_{\mathbb{S}^1}))^\vee.$$

Furthermore, establish formula (5.5): $\nabla_{\dot c}\alpha(\cdot)(x) = \nabla^g_{\dot c^\wedge(\cdot,x)}\alpha^\wedge(\cdot,x)$, for all $x \in \mathbb{S}^1$.

5.1.3 Continue Example 5.10 and compute for $C^\infty(\mathbb{S}^1, H)$ with the L^2-metric an explicit form of the geodesic equation $\nabla_{\dot c}\dot c = 0$. Deduce again that a geodesic $c\colon J \to C^\infty(\mathbb{S}^1, H)$ is given for each $\theta \in \mathbb{S}^1$ by the affine linear map $c(t)(\theta) = tc'(0)(\theta) + c(0)(\theta)$.

5.1.4 Let (M, g) be a Riemannian manifold with a metric spray and metric derivative ∇. Work in a local chart to establish the identity $X.g(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$, (5.3), for the covariant derivative along a smooth map $F\colon N \to M$.

5.2 Shape Analysis via the Square Root Velocity Transform

We return to the announced application of the L^2-metric to shape analysis. As we have seen in the introduction, we seek to construct a shape space together with a Riemannian structure which will allow us to compare its elements using geodesics and geodesic distance. We begin by defining the necessary spaces and metrics.

5.12 Define the *pre-shape space* of closed curves

$$\mathcal{P} := \operatorname{Imm}(\mathbb{S}^1, \mathbb{R}^2) = \{c \in C^\infty(\mathbb{S}^1, \mathbb{R}^2) \mid \dot c(t) \neq 0, \text{ for all } t \in [0,1]\}.$$

In Exercise 2.1.3 and Example 4.6 we have seen that $\mathcal{P}$ is an open subset of $C^\infty(\mathbb{S}^1, \mathbb{R}^2)$ which becomes a weak Riemannian manifold with respect to the Riemannian metric

$$g_c(f,g) = \int_{\mathbb{S}^1} \langle f(\theta), g(\theta)\rangle \|\dot c(\theta)\| \mathrm{d}\theta.$$

We are actually interested in the images of elements in $\mathcal{P}$ and want to identify all curves which yield the same image up to a reparametrisation. To model the reparametrisation, consider the group

$$\operatorname{Diff}(\mathbb{S}^1) := \{\varphi \in C^\infty(\mathbb{S}^1, \mathbb{S}^1) \mid \varphi \text{ is bijective with smooth inverse}\}$$

of diffeomorphisms of $\mathbb{S}^1$. The group acts on $\mathcal{P}$ via reparametrisation, that is,

$$\gamma\colon \operatorname{Diff}(\mathbb{S}^1) \times \mathcal{P} \to \mathcal{P}, \quad (\varphi, c) \mapsto c \circ \varphi$$

is a Lie group action; see Example 3.5. We can thus define the *shape space* $\mathcal{S} = \mathcal{P}/\operatorname{Diff}(\mathbb{S}^1)$ as the quotient of $\mathcal{P}$ with respect to the Lie group action.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

One can show that $\mathcal{S}$ is almost a manifold[1] and since the Riemannian metric is invariant under the reparametrisation action, the Riemannian metric induces a Riemannian metric on $\mathcal{S}$. Unfortunately, Michor and Mumford (2006, 3.10) have shown that g has vanishing geodesic distance (see also Example 4.14), whence any attempt to compare shapes by computing their geodesic distance has to fail.

The defect of the weak Riemannian metric can be solved by incorporating derivatives in the definition of the Riemannian metric. This leads to the notion of a family of metric called H^1-metrics (the name indicates that the associated strong Riemannian manifold consists of Sobolev H^1-functions).

5.13 (An elastic inner product, Mio et al., 2007) Let $c \in \operatorname{Imm}(\mathbb{S}^1,\mathbb{R}^2)$ with $\dot{c} = (c,c')$. Then we define $u_c(\theta) := c'(\theta)/\|\dot{c}(\theta)\|$ and the arc length derivative $D_{c,\theta}(h) = h'/\|\dot{c}\|$. We define an inner product on $T_c \operatorname{Imm}(\mathbb{S}^1,\mathbb{R}^2) = C^\infty(\mathbb{S}^1,\mathbb{R}^2)$:

$$G_c(h,k) := \int_{\mathbb{S}^1} \frac{1}{4}\langle D_{c,\theta}h, u_c\rangle\langle D_{c,\theta}k, u_c\rangle + \langle D_{c,\theta}h - u_c\langle D_{c,\theta}h, u_c\rangle, D_{c,\theta}k - u_c\langle D_{c,\theta}k, u_c\rangle\rangle\|\dot{c}\|\mathrm{d}\theta. \tag{5.9}$$

This inner product is called *elastic inner product* as the first term measures stretching in the direction of c, while the second term measures bending of the curve c. Note that due to its construction, the elastic inner products are invariant under the reparametrisation action of $\operatorname{Diff}(\mathbb{S}^1)$ on $\mathcal{P}$.

It is not hard to see that these inner products yield a weak Riemannian metric on the pre-shape space $\mathcal{P}$. We will derive this only for a smaller submanifold using the so-called *square root velocity transform (SRVT)*:

5.14 Definition Define the mappings

$$\begin{aligned}
\mathcal{R}\colon \mathcal{P} = \operatorname{Imm}(\mathbb{S}^1,\mathbb{R}^2) &\to \{q \in C^\infty(\mathbb{S}^1,\mathbb{R}^2) \mid q(\theta) \neq 0 \text{ for all } \theta \in \mathbb{S}^1\} \\
&= C^\infty(\mathbb{S}^1,\mathbb{R}^2 \setminus \{0\}) \\
c &\mapsto \mathcal{R}(c)(t) := c'(t)/\sqrt{\|\dot{c}(t)\|}, \\
\mathcal{R}^{-1}\colon C^\infty(\mathbb{S}^1,\mathbb{R}^2 \setminus \{0\}) &\to \mathcal{P}, \\
q &\mapsto \left((\cos(t),\sin(t)) \mapsto \int_0^t q(\cos(s),\sin(s)) \cdot \|q(\cos(s),\sin(s)\|\mathrm{d}s\right).
\end{aligned}$$

[1] There exist singularities, turning $\mathcal{S}$ into an orbifold. See Example 6.4. The existence of singularities is usually ignored in shape analysis, as an open dense subset of $\mathcal{S}$ is a manifold such that the projection restricts on this set to a submersion.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

We call $\mathcal{R}$ the *square root velocity transform* (SRVT).

Using the idea that $\mathcal{P} \subseteq C^\infty(\mathbb{S}^1,\mathbb{R}^2)$ and that pushforwards of smooth mappings are smooth maps between canonical manifolds of mappings (see Chapter 2), it is not hard to see that $\mathcal{R}$ and $\mathcal{R}^{-1}$ are smooth. Moreover, $\mathcal{R}\circ\mathcal{R}^{-1}(q) = q$, but since $\mathcal{R}$ involves differentiation, we lose information on the starting point of the curve and have $\mathcal{R}^{-1}\circ\mathcal{R}(c) = c$ if and only if $c(\cos(0),\sin(0)) = 0$, that is, the curve starts at the origin. As we are interested in shapes, it will be irrelevant as to where in $\mathbb{R}^2$ the shape is located. In other words, we can restrict to the submanifold

$$\mathcal{P}_* \coloneqq \{c \in \mathcal{P} \mid c(\cos(0),\sin(0)) = 0\}$$

of all immersions starting at the origin. We will see in Exercise 5.2.1 that the SRVT induces a diffeomorphism between $\mathcal{P}_*$ and the manifold $C^\infty(\mathbb{S}^1,\mathbb{R}^2\setminus\{0\})$. Recall now the notion of pullback of a Riemannian metric.

5.15 Definition Let (M,g) be a (weak) Riemannian manifold and $\varphi\colon N \to M$ be an immersion. Then N can be made a (weak) Riemannian manifold with respect to the *pullback metric* defined via

$$(\varphi^* g)_m(v,w) \coloneqq g_{\varphi(m)}(T_m\varphi(v), T_m\varphi(w)).$$

5.16 Example If $U \subseteq M$ and (M,g) is a weak Riemannian manifold, then the inclusion $\iota\colon U \to M$ is an immersion and the restriction of g to U coincides with the pullback metric obtained from ι. In particular, an open subset of a weak Riemannian manifold, such as $C^\infty(\mathbb{S}^1,\mathbb{R}^2\setminus\{0\}) \subseteq C^\infty(\mathbb{S}^1,\mathbb{R}^2)$ (with the weak L^2-metric), becomes a weak Riemannian manifold by restriction.

5.17 Proposition *The pullback metric of the non-invariant L^2-metric on $C^\infty(\mathbb{S}^1,\mathbb{R}^2)$ from Example 4.3 with respect to the SRVT $\mathcal{R}$ is the* elastic metric *described by (5.9) on each tangent space of $\mathcal{P}_*$.*[2]

Proof By construction, the square root velocity transform $\mathcal{R}$ is the composition of the derivative operator $D\colon \operatorname{Imm}(\mathbb{S}^1,\mathbb{R}^2) \to C^\infty(\mathbb{S}^2,\mathbb{R}^2\setminus\{0\})$, $q \mapsto (\theta \mapsto q'(\theta) = T_\theta c(1))$, and the scaling map $\operatorname{sc}\colon C^\infty(\mathbb{S}^1,\mathbb{R}\setminus\{0\}) \to C^\infty(\mathbb{S}^1,\mathbb{R}\setminus\{0\})$, $f \mapsto f/\sqrt{\|f\|}$. Thus by the chain rule we have $T_c\mathcal{R} = T\operatorname{sc}\circ T_c D$ and to arrive at the desired formula, we have to compute the derivatives of these two mappings. To this end, we exploit that $TC^\infty(\mathbb{S}^1,\mathbb{R}^2) = C^\infty(\mathbb{S}^1,T\mathbb{R}^2) = C^\infty(\mathbb{S}^1,\mathbb{R}^2\times\mathbb{R}^2)$ and since $\operatorname{Imm}(\mathbb{S}^1,\mathbb{R}^2) \subseteq C^\infty(\mathbb{S}^1,\mathbb{R}^2)$, we can similarly identify the tangent bundle of the immersions. Now arguments as in Exercise 1.2.2(c) show that the differential operator D is continuous linear, whence for an element (c,V)

[2] The statement remains valid if we replace $\mathbb{R}^2$ by an arbitrary Hilbert space of dimension ≥ 2.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

in $T_c C^\infty(\mathbb{S}^1,\mathbb{R}^2) = \{(c,V) \in C^\infty(\mathbb{S}^1,\mathbb{R}^2)^2\}$, we obtain $T_c D(V) = V'$. In Exercise 5.2.2 we shall show that the derivative of the scaling map is

$$T_q \mathrm{sc}(Z) = \frac{Z}{\sqrt{\|q\|}} - \frac{1}{2\sqrt{\|q\|}^5}\langle Z,q\rangle q. \tag{5.10}$$

To obtain the derivative of the SRVT at (c,V), we simply have to replace q with c' and Z with V'. This yields the following formula:

$$\begin{aligned}(\mathcal{R}^*\langle\cdot,\cdot\rangle_{L^2})_c(V,W) &= \langle T_c\mathcal{R}(V), T_c\mathcal{R}(W)\rangle_{L^2} = \int_{\mathbb{S}^1}\langle T_{c'}\mathrm{sc}(V')(\theta), T_{c'}\mathrm{sc}(W)(\theta)\rangle \mathrm{d}\theta \\ &= \int_{\mathbb{S}^1}\left\langle \frac{V'}{\sqrt{\|c'\|}} - \frac{1}{2\sqrt{\|c'\|}^5}\langle V',c'\rangle c', \frac{W'}{\sqrt{\|c'\|}} - \frac{1}{2\sqrt{\|c'\|}^5}\langle W',c'\rangle c'\right\rangle(\theta)\mathrm{d}\theta.\end{aligned} \tag{5.11}$$

Since the inner product is bilinear, we can factor out terms of the form $\|c'\|$ and replace V', W' and c' with their rescaled versions (see 5.13). Now an easy but tedious computation shows that the pullback metric (5.11) coincides with the elastic metric (5.9). □

Hence the elastic metric can be understood by studying the L^2-metric on the manifold $C^\infty(\mathbb{S}^1,\mathbb{R}^2\setminus\{0\})$. However, as this is just an open subset of the (weak) Riemannian manifold $(C^\infty(\mathbb{S}^1,\mathbb{R}^2),\langle\cdot,\cdot\rangle^{L^2})$ we already know the spray, connector, covariant derivative and geodesics of this metric from our discussion of the L^2-metric. It is important to observe that the elastic metric pulls back to the non-invariant L^2-metric. For the non-invariant L^2-metric, the geodesic distance does not vanish (compare this with the invariant L^2-metric, Example 4.6) and the pullback metric (i.e. the elastic metric) is invariant under reparametrisation.

A natural extension for the vector space valued shape spaces discussed in Proposition 5.17 is Lie group valued shape spaces. Here a shape is (up to quotienting out the reparametrisation action) an element of the loop group $C^\infty(\mathbb{S}^1,G)$. If G is a Hilbert Lie group, the construction of the square root velocity transform can be adapted to this more general setting by using the logarithmic derivative of Lie group valued mappings. For details, we refer to Exercise 5.2.4. Similar techniques have been used in Celledoni et al. (2018) for shape analysis on homogeneous spaces.

5.18 Example (Sample application: Motion capturing; see e.g. Celledoni et al., 2016) Assume that we have motion capturing data of, for example, a human walking, given by a number of time-dependent datapoints in $\mathbb{R}^3$. The associated virtual character is then modelled as a skeleton for which these datapoints represent positions of certain parts. Shifting the focus from the position to their relative position, we can interpret the positions of every part of the skeleton as

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

an angle between neighbouring parts. Now angles in $\mathbb{R}^3$ can be identified with rotations, that is, elements in the Lie group SO(3) of rotations of $\mathbb{R}^3$. Thus if we do not impose constraints on the allowed angles (which can lead to unnatural movements, but is generally fine when working with real motion-capturing data), we can think of motion-capturing data as a smooth curve with values in a product of copies of SO(3) (the number depends on the number of data points which move relative to each other). In Celledoni et al. (2016) numerical algorithms for automatic interpolation and transformation of motion-capturing data have been constructed which exploit the above point of view.

5.19 Remark Note that the geodesic distance of the L^2-metric on $C^\infty(\mathbb{S}^1, H)$ does not vanish and is indeed positive for any two shapes which are not equal (where H is a Hilbert space). Indeed the geodesic distance just coincides with the L^2-distance of the curves. The situation gets more complicated in the space $C^\infty(\mathbb{S}^1, H \setminus \{0\})$ and, in particular, for $H = \mathbb{R}^2$ (which is not simply connected), so an element γ for which 0 is contained in a bounded connected component of $\mathbb{R}^2 \setminus \gamma(\mathbb{S}^1)$ cannot be connected by a continuous curve in $C^\infty(\mathbb{S}^1, \mathbb{R}^2 \setminus \{0\})$ to an element for which 0 is not contained in such a component.

While this does not happen for open shapes (i.e. elements of $\operatorname{Imm}([0,1], \mathbb{R}^2)$), the following problem is even more significant when it comes to applications in shape analysis: As geodesics are not allowed to pass through 0, the L^2-distance of two functions is *not* the geodesic distance even if there exist continuous paths between them. If the linear interpolation between two points $c(\theta)$ and $d(\theta)$ passes through 0, then the geodesic from the ambient space $C^\infty(\mathbb{S}^1, \mathbb{R}^2)$ does not exist in the smaller space. Instead, the geodesic distance then needs to be computed using curves which 'move one shape around the hole'. This leads to a geodesic distance, which is strictly larger than the L^2-distance. In numerical applications this is often just plainly ignored.

That the geodesic distance on $C^\infty(\mathbb{S}^1, \mathbb{R}^2 \setminus \{0\})$ diverges from the easily computed L^2-distance is a consequence of the space being incomplete. Indeed since we have 'drilled a hole' by excluding a point, geodesics passing through that point can only exist up to the time they enter. As the L^2-distance moves points on the image of functions along straight lines interpolating between them, the L^2-distance fails to give the geodesic distance for points lying on opposite sides of the excluded point.

Exercises

5.2.1 Establish some properties of the square root velocity transform $\mathcal{R}$, Definition 5.14.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Hint: Use the substitution rule for integrals on submanifolds; in the case at hand, this works just as in usual interval cases.
Show that:

(a) $\mathcal{R}$ and $\mathcal{R}^{-1}$ are smooth maps with $\mathcal{R} \circ \mathcal{R}^{-1} = \mathrm{id}_{C^\infty(\mathbb{S}^1,\mathbb{R}^2\setminus\{0\})}$.
(b) $\mathcal{P}_* = \{c \in \mathcal{P} \mid c(\cos(0),\sin(0)) = 0\}$ is a closed submanifold of $\mathcal{P}$.
(c) $\mathcal{R}$ and $\mathcal{R}^{-1}$ induce a diffeomorphism between the manifolds $\mathcal{P}_*$ and $C^\infty(\mathbb{S}^1,\mathbb{R}^2 \setminus \{0\})$.
(d) If $\varphi \in \mathrm{Diff}(\mathbb{S}^1)$ with $\varphi'(\theta) > 0$, for all $\theta \in \mathbb{S}^1$, then $\mathcal{R}(c \circ \varphi) = \sqrt{\varphi'} \cdot \mathcal{R}(c) \circ \varphi$.

5.2.2 Let $(E, \|\cdot\|)$ be a Hilbert space with $\|v\|^2 = \langle v, v\rangle$ and K a compact manifold. Prove that the scaling map $\mathrm{sc}\colon C^\infty(K,E\setminus\{0\}) \to C^\infty(K,E\setminus\{0\})$, $q \mapsto q/\sqrt{\|q\|}$ is smooth with the tangent map given by the formula (5.10):

$$T_q\mathrm{sc}(W) = \frac{W}{\sqrt{\|q\|}} - \frac{1}{2\sqrt{\|q\|}^5}\langle W, q\rangle q.$$

Hint: Use the exponential law together with the canonical identification of the tangent bundles.

5.2.3 Show that the elastic metric (5.9) is invariant under reparametrisations with elements φ in $\mathrm{Diff}(\mathbb{S}^1)$ which satisfy $T_\theta\varphi(1) > 0$, for all $\theta \in \mathbb{S}^1$.

5.2.4 Let G be a Hilbert Lie group, that is, $\mathbf{L}(G)$ is a Hilbert space with inner product $\langle\cdot,\cdot\rangle$. Then we define a square root velocity transform on the subset of immersions of the loop group

$$\mathcal{R}\colon \mathrm{Imm}(\mathbb{S}^1,G) \to C^\infty(\mathbb{S}^1,\mathbf{L}(G)\setminus\{0\}), \quad c \mapsto \delta^r(c)/\sqrt{\|\delta^r(c)\|},$$

where δ^r is the right logarithmic derivative.

(a) Show that $\mathcal{R}$ is a smooth diffeomorphism (what is its inverse?).
(b) Compute a formula for the pullback of the L^2-metric on $C^\infty(\mathbb{S}^1,\mathbf{L}(G)\setminus\{0\})$. This metric is known as the elastic metric on the Lie group valued immersions and it can be used in computer animation and motion-capturing applications. See Celledoni et al. (2016) for more information.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

6

Connecting Finite-Dimensional, Infinite-Dimensional and Higher Geometry

In this chapter we will highlight an interesting connection between finite- and infinite-dimensional differential geometry. To this end, we shall consider in §6.2 elements from 'higher geometry', so-called Lie groupoids. The moniker higher geometry stems from the fact that in the language of category theory, these objects form higher categories. We shall not explore higher categories or their connection to differential geometry in this book (but the reader might consult Meyer and Zhu (2015) or the general introduction by Baez (1997)). In previous chapters we have discussed how finite-dimensional manifolds and geometric structures give rise to infinite-dimensional structures such as Lie groups (e.g. the diffeomorphism and groups of gauge transformations) and Riemannian metrics (such as the L^2-metric from shape analysis). While we have seen that every manifold determines an (in general, infinite-dimensional) group of diffeomorphisms, we turn this observation now on its head and ask:

Can we recognise the underlying finite-dimensional geometric structure from the infinite-dimensional object?

6.1 Diffeomorphism Groups Determine Their Manifolds

Let us examine this question for the diffeomorphism group. We have already seen that for every compact manifold we can associate the infinite-dimensional diffeomorphism group. Conversely, Takens (1979) has shown that the diffeomorphism group identifies (up to diffeomorphism) the underlying manifold. Namely, we have the following theorem (which we cite here from the much more general statement of Filipkiewicz (1982)):

6.1 Theorem (Takens (1979)/Filipkiewicz (1982)/Banyaga (1988)/Rubin (1989)) *If M, N are smooth compact, connected manifolds such that*

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$\phi\colon \mathrm{Diff}(M) \to \mathrm{Diff}(N)$ is a group isomorphism then there exists a diffeomorphism $\phi\colon M \to N$ such that $\Phi(\gamma) = \phi \circ \gamma \circ \phi^{-1}$.

A full proof of Theorem 6.1 would lead us too far astray, but it is possible to highlight certain aspects of the proof which are of special interest to us with regard to the question of whether finite-dimensional objects can be recognised from their associated infinite-dimensional objects. Before we begin, let us recall two concepts for diffeomorphism groups.

6.2 Definition Let M be a smooth and compact manifold.

- For $x_0 \in M$, the *stabiliser* is defined as

$$S_{x_0}\,\mathrm{Diff}(M) \coloneqq \{h \in \mathrm{Diff}(M) \mid h(x_0) = x_0\}.$$

- The group $\mathrm{Diff}(M)$ *acts n-transitive* on the manifold M for $n \in \mathbb{N}$ if for any two sets $\{x_1,\dots,x_n\},\{y_1,\dots,y_n\}$ of non-repeating points in M, there is $h \in \mathrm{Diff}(M)$ with $h(x_i) = y_i, i \in \{1,\dots,n\}$.

We shall show how a group isomorphism of diffeomorphism groups mapping stabilisers to each other induces a diffeomorphism on the base manifold. The first step towards this is Banyaga (1988, Lemma 1) (also compare Rybicki, 1995).

6.3 Lemma *Let M, N be two connected smooth manifolds and $\phi\colon \mathrm{Diff}(M) \to \mathrm{Diff}(N)$ a group isomorphism such that the following holds:*

(a) *For $K \in \{M,N\}$, $\mathrm{Diff}(K)$ acts n-transitively for $n \in \{1,2\}$, and*
(b) *for each $x_0 \in M$ there exists $y_0 \in N$ such that $\phi(S_{x_0}\,\mathrm{Diff}(M)) = S_{y_0}\,\mathrm{Diff}(N)$.*

Then there is a unique homeomorphism $\omega\colon M \to N$ such that $\phi(f) = \omega f \omega^{-1}$.

Proof *Step 1: Construction of the homeomorphism ω.* Fix a pair of points x_0 and y_0 as in condition (b). Since $\mathrm{Diff}(M)$ and $\mathrm{Diff}(N)$ are 1-transitive, we see that

$$\mathrm{ev}_{x_0}\colon \mathrm{Diff}(M) \to M, \quad h \mapsto h(x_0) \text{ and } \mathrm{ev}_{y_0}\colon \mathrm{Diff}(N) \to N, \quad g \mapsto g(y_0)$$

are surjective with $\mathrm{ev}_{x_0}^{-1}(x_0) = S_{x_0}\,\mathrm{Diff}(M)$ (and similarly for y_0). We can thus choose for every $x \in M$ an (in general) non-unique diffeomorphism $h_x \in \mathrm{Diff}(M)$ such that $h_x(x_0) = x$. Now if $\tilde{h}$ is another diffeomorphism such that $\tilde{h}(x_0) = x$ we see that $\tilde{h}^{-1}h_x(x_0) = x_0$, and so $\tilde{h}^{-1}h_x \in S_{x_0}\,\mathrm{Diff}(M)$. As by assumption, $\phi(S_{x_0}\,\mathrm{Diff}(M)) = S_{y_0}\,\mathrm{Diff}(N)$, this implies $\phi(\tilde{h}^{-1}h_x) =$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$\phi(\tilde{h}^{-1})\phi(h_x) \in S_{y_0}\operatorname{Diff}(N)$. Indeed, the value is independent of the choice of h_x, whence we obtain a well-defined map

$$\omega\colon M \to N, \quad x \mapsto \operatorname{ev}_{x_0}(\phi(h_x)) = \phi(h_x)(y_0)$$

as this mapping does not depend on the choice of h_x. Again by the above, the map ω is a bijection which is even a homeomorphism (details will be checked in Exercise 6.1.1).

Step 2: ω induces ϕ. Let $y \in N$ and $h \in \operatorname{Diff}(N)$ with $h(y_0) = y$ and $x = \phi^{-1}(h)(x_0)$. By construction we have $\omega(x) = y$. If $f \in \operatorname{Diff}(M)$ we pick $g \in \operatorname{Diff}(M)$ with $g(x_0) = f(x)$. Then $f^{-1}(g(x_0)) = x = \phi^{-1}(h)(x_0)$ and thus $g^{-1}f\phi^{-1}(h) \in S_{x_0}\operatorname{Diff}(M)$. We deduce that $(\phi(g))^{-1} \circ \phi(f) \circ h \in S_{y_0}\operatorname{Diff}(N)$ or in other words $\phi(f) \circ h(y_0) = \phi(g)(y_0)$. Now $h(y_0) = y = \omega(x)$ and $\phi(g)(y_0) = \omega(f(x))$ (as $g(x_0) = f(x)$). We deduce that

$$\phi(f)(\omega(x)) = \omega(f(x)), \quad \phi(f) \circ \omega = \omega \circ f,$$

and using that ω is bijective, this yields $\phi(f) = \omega \circ f \circ \omega^{-1}$.

Step 3: ω is unique. Assume that there is another homeomorphism $\tilde{\omega}\colon M \to N$ inducing ϕ. Then

$$\omega \circ f \circ \omega^{-1} = \phi(f) = \tilde{\omega} \circ f \circ \tilde{\omega}, \quad \text{for all } f \in \operatorname{Diff}(M).$$

In other words we have for $\rho := \tilde{\omega}^{-1} \circ \omega$ that $\rho \circ f \circ \rho^{-1} = f$, for all $f \in \operatorname{Diff}(M)$. Arguing by contradiction we assume that $\rho \neq \mathrm{id}_M$. Then there exists $x \in M$ with $y = \rho(x) \neq x$. Pick $z \in M \setminus \{x, y\}$. Now $\operatorname{Diff}(M)$ is 2-transitive, whence there is $f \in \operatorname{Diff}(M)$ with $f(x) = x$ and $f(y) = z$. We see that

$$\rho \circ f \circ \rho^{-1}(y) = \rho(f(x)) = \rho(x) = y \neq z = f(y).$$

However, this contradicts $\rho \circ f \circ \rho^{-1} = f$, hence we must have $\tilde{\omega} = \omega$. □

Sketch of the proof of Theorem 6.1 Every group isomorphism $\phi\colon \operatorname{Diff}(M) \to \operatorname{Diff}(N)$ satisfies condition (b) in the statement of Lemma 6.3 (this is far from trivial; see Filipkiewicz, 1982). Moreover, the group of smooth diffeomorphisms acts n-transitively for every $n \in \mathbb{N}$ if $\dim M > 1$; see Michor and Vizman (1994). Thus we can apply Lemma 6.3 to obtain a homeomorphism $\omega\colon M \to N$.

We prove that ω is a diffeomorphism under the assumption that $\phi\colon \operatorname{Diff}(M) \to \operatorname{Diff}(N)$ is a Lie group isomorphism. Note that the statement of the theorem is much stronger as, a priori, ϕ need not even be continuous. However, in this case one needs a deep result on Lie group actions on manifolds; see Filipkiewicz (1982, Step 3 on p. 173). A posteriori this implies that any group isomorphism $\operatorname{Diff}(M) \to \operatorname{Diff}(N)$ is already a Lie group isomorphism.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

So let us assume that ϕ is a Lie group isomorphism and let us study the composition $\omega \circ \mathrm{ev}_{x_0} \colon \mathrm{Diff}(M) \to N$. By Step 2 of the proof of Lemma 6.3 we can rewrite this as

$$\omega \circ \mathrm{ev}_{x_0}(h) = \mathrm{ev}_{y_0}(\phi(h)) \quad (\text{and conversely } \omega^{-1} \circ \mathrm{ev}_{y_0} = \mathrm{ev}_{x_0} \circ \phi^{-1}).$$

Now Exercise 2.3.3(a) shows that ev_{x_0} and ev_{y_0} are smooth surjective submersions. Hence the smoothness of the right-hand side together with Exercise 1.7.5 shows that ω and ω^{-1} are smooth. □

We have seen that diffeomorphism groups determine (up to diffeomorphism) their underlying manifold uniquely. In the next section we shall discuss objects, so-called Lie groupoids, which can be used to describe many finite-dimensional geometric structures and which admit a similar connection to infinite-dimensional groups.

Exercises

6.1.1 We are working in the setting of Lemma 6.3 and let $\phi \colon \mathrm{Diff}(M) \to \mathrm{Diff}(N)$ be a group isomorphism which maps the stabiliser $S_{x_0}\mathrm{Diff}(M)$ to the stabiliser $S_{y_0}\mathrm{Diff}(N)$.

(a) Show that the mapping $\omega \colon M \to N$, $x \mapsto \phi(h_x)(y_0)$ is a bijection (where $h_x \in \mathrm{Diff}(M)$ with $h_x(x_0) = x$).

(b) It is well known that $\mathrm{Diff}(M)$ satisfies the following condition: *For any non-empty connected $U \subsetneq M$ and $x \in U$, there exists $h \in \mathrm{Diff}(M) \setminus \{\mathrm{id}_M\}$ such that $\overline{\{y \in M \mid h(x) \neq x\}} \subseteq U$ and x is contained in the interior of $\overline{\{y \in M \mid h(x) \neq x\}}$.*
Let $f \in \mathrm{Diff}(M)$ and define $\mathrm{Fix}(f) := \{x \in M \mid f(x) = x\}$. Show that the set $\mathcal{B} := \{M \setminus \mathrm{Fix}(f) \mid f \in \mathrm{Diff}(M)\}$ is a basis for the topology on M.

(c) Show that $\mathrm{Fix}(\phi(g)) = \omega(\mathrm{Fix}(g))$ holds and conclude that ω is a homeomorphism.

6.2 Lie Groupoids and Their Bisections

In this section we consider a generalisation of Lie groups called Lie groupoids. These objects allow one to treat constructions in differential geometry as differentiable objects. For example, quotients of manifolds modulo Lie group actions may fail to be manifolds. However, one can encode them using suitable

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Lie groupoids. We motivate the construction with the following example of an ill-behaved quotient.

6.4 Example In Chapter 5, we studied shape spaces which arise as quotients of manifolds of mappings modulo an action of the diffeomorphism group. Namely, we considered the canonical action of the Lie group $\mathrm{Diff}(\mathbb{S}^1)$ on the open submanifold $\mathrm{Imm}(\mathbb{S}^1,\mathbb{R}^2) \subseteq C^\infty(\mathbb{S}^1,\mathbb{R}^2)$ (see Example 3.5 and Lemma 2.6) via precomposition

$$p\colon \mathrm{Imm}(\mathbb{S}^1,\mathbb{R}^2)\times \mathrm{Diff}(\mathbb{S}^1) \to \mathrm{Imm}(\mathbb{S}^1,\mathbb{R}^2), \quad (f,\varphi)\mapsto f\circ\varphi.$$

The shape space $\mathcal{S} := \mathrm{Imm}(\mathbb{S}^1,\mathbb{R}^2)/\mathrm{Diff}(\mathbb{S}^1)$ is then the quotient modulo the action. It inherits a natural topology which, however, does not turn $\mathcal{S}$ into a manifold. As the action p is not free, the quotient has singular points in which one fails to obtain charts. An example for such a point is the image of the immersion of the circle into $\mathbb{R}^2$ given by the map tracing out the circle in 'double speed':

$$c\colon \mathbb{S}^1 \to \mathbb{R}^2, \quad e^{i\theta}\mapsto e^{i2\theta}.$$

Since c traces the circle twice we see that $p(c,\varphi) = c\circ\varphi = c$ for the diffeomorphism $\varphi\colon \mathbb{S}^1 \to \mathbb{S}^1$, $e^{i\theta}\mapsto e^{i(\theta+\pi)}$, and we deduce that the immersion c has a non-trivial stabiliser under the action ρ. Thus, in particular, the quotient $\mathcal{S}$ is not a manifold (though every singularity is mild in the sense that it is generated by a finite group, i.e. one obtains an infinite-dimensional orbifold; see Michor, 2020, Section 7.3). As manifolds are the basic setting for differential geometry, one needs to pass to the subset generated by the free immersions (i.e. those immersions with trivial stabiliser). These indeed form a dense open subset which is a manifold.

The last example exhibits that quotients of Lie group actions will, in general, not be manifolds. While in this special example one could still say a lot about the structure of the quotient, it shows that quotient constructions with manifolds are, in general, very badly behaved (not only in infinite dimensions). Thus we would like to avoid quotients and obtain an object which contains the same information as the quotient: a (Lie) groupoid. There are many literature accounts for the basic theory of (finite-dimensional) Lie groupoids such as Mackenzie (2005) and Meinrencken (2017). While the finite-dimensional examples will be most important for us as they describe geometric constructions, the concept of a Lie groupoid can also be formulated in the infinite-dimensional context (as the notion of submersion makes sense in this setting).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

6.5 Definition (Groupoid) Let G, M be two sets with surjective maps $\mathbf{s}, \mathbf{t}\colon G \to M$ (*source* and *target*) and a partial multiplication $\mathbf{m}\colon G \times G \supseteq (\mathbf{s}, \mathbf{t})^{-1}(M \times M) \to G$, $(a, b) \mapsto ab$ which satisfies:

(a) $\mathbf{s}(ab) = \mathbf{s}(b)$ and $\mathbf{t}(ab) = \mathbf{t}(a)$, and $(ab)c = a(bc)$;
(b) *identity section* $\mathbf{1}\colon M \to G$ with $\mathbf{1}(\mathbf{t}(g))g = g$ and $g\,\mathbf{1}(\mathbf{s}(g)) = g$ for all $g \in G$;
(c) *inverses* for all $g \in G$ there is $g^{-1} \in G$ with $g^{-1}g = \mathbf{1}(\mathbf{s}(g))$ and $gg^{-1} = \mathbf{1}(\mathbf{t}(g))$.

We call G (or $\mathcal{G} = (G \rightrightarrows M)$) a *groupoid* and the set M is called the *set of units*. If G, M are smooth manifolds, such that the structure maps $\mathbf{s}, \mathbf{t}$ are smooth submersions and $\mathbf{m}, \mathbf{1}$ and the inversion map $\mathbf{i}\colon G \to G, \mathbf{i}(g) = g^{-1}$ are smooth maps, we say that $\mathcal{G} = (G \rightrightarrows M)$ is a *Lie groupoid*.

Standard Notation for Lie Groupoids Throughout this section (if nothing else is said), we write $\mathcal{G} = (G \rightrightarrows M)$ for a Lie groupoid with structure maps $\mathbf{1}$, $\mathbf{s}$, $\mathbf{t}$, $\mathbf{m}$, $\mathbf{i}$ as in the definition of a groupoid.

6.6 Remark In this book manifolds are required to be Hausdorff. Thus Definition 6.5 excludes by design Lie groupoids $\mathcal{G} = (G \rightrightarrows M)$ whose space of arrows G is not Hausdorff. A broad (and important) class of Lie groupoids with non-Hausdorff space of arrows are the so-called foliation groupoids, arising from the treatment of foliations in a groupoid framework; see Moerdijk and Mrčun (2003). In principle many of the results presented here are also valid in the non-Hausdorff setting; see, for example, Rybicki (2002).

A useful mental image to keep in mind is to picture the units of the groupoid as dots connected by arrows which represent the elements of the groupoid which are not units. This is illustrated in Figure 6.1. Obviously two arrows can then be composed only if one of them ends where the other starts. This picture also immediately shows in which way a groupoid generalises the concept of a group: It can possess more units and its elements are not necessarily composable.

6.7 Definition For $G \rightrightarrows M$ a groupoid, and $a \in M$ a unit, we consider the fibres $\mathbf{s}^{-1}(a)$ and $\mathbf{t}^{-1}(a)$ of all arrows starting, resp. ending at a. The intersection $G_a := \mathbf{s}^{-1}(a) \cap \mathbf{t}^{-1}(a)$ forms a group, called the *vertex group* at a of the groupoid.

If $(G \rightrightarrows M)$ is a Lie groupoid, $\mathbf{s}^{-1}(a)$ and $\mathbf{t}^{-1}(a)$ are submanifolds of G by Corollary 1.59. A natural question is then whether the vertex groups G_a inherit a Lie group structure. In general (for arbitrary infinite-dimensional Lie groupoids), this question is still open as the finite-dimensional argument

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Figure 6.1 Picturing groups and groupoids. In the right picture we suppressed all arrows between the two nodes with looping arrows. As arrows tracing a path from one node to the other can always be composed, a picture of all groupoid elements would also need to represent these arrows.

(Mackenzie, 2005, Corollary 1.4.11) establishing the Lie group structure breaks down in infinite dimensions. However, it has recently been proven in Beltiţă et al. (2019) that if G, M are Banach manifolds, then the vertex groups of $G \rightrightarrows M$ are (Banach) Lie groups. Before we finally give examples for Lie groupoids, let us first define groups which will play roles similar to the diffeomorphism group in the previous section in relation to a manifold.

6.8 Definition The *group of bisections* $\mathrm{Bis}(\mathcal{G})$ of $\mathcal{G}$ is given as the set of smooth maps $\sigma\colon M \to G$ such that $\mathbf{s}\circ\sigma = \mathrm{id}_M$ and $\mathbf{t}\circ\sigma\colon M \to M$ is a diffeomorphism. The group structure is given by the product

$$(\sigma \star \tau)(x) := \sigma((\mathbf{t}\circ\tau)(x))\tau(x) \text{ for } x \in M. \tag{6.1}$$

The identity section $\mathbf{1}\colon M \to G$ is the neutral element and the inverse of σ is

$$\sigma^{-1}(x) := \mathbf{i}(\sigma((\mathbf{t}\circ\sigma)^{-1}(x))) \text{ for } x \in M. \tag{6.2}$$

The definition of bisection is not symmetric with respect to source and target. This lack of symmetry can be avoided by defining a bisection as a set (see Mackenzie, 2005, p. 23). However, this point of view does not fit well into the function space perspective we take. Thus we shall stick with the asymmetric definition.

Lie groupoids simultaneously generalise Lie groups and (differentiable) equivalence relations. To emphasise this, we recall the following standard examples (see Mackenzie, 2005).

6.9 Example In the following, we denote by $\{\bullet\}$ the one-point manifold.

(a) Let G be a Lie group. Then the Lie group structure yields a Lie groupoid $G \rightrightarrows \{\bullet\}$, that is, a Lie group is a Lie groupoid and conversely every Lie groupoid whose set of units contains only one element is a Lie group. In this case $\mathrm{Bis}(G \rightrightarrows \{\bullet\}) = G$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(b) Let $\pi\colon M \to N$ be a submersion. Then the fibre product of M with itself gives rise to a Lie groupoid $M \times_N M \rightrightarrows M$. Its source and target maps are given as $\mathbf{s} = \mathrm{pr}_2$ and $\mathbf{t} = \mathrm{pr}_1$. Multiplication is then given by concatenation $(m,n)\cdot(n,k) := (m,k)$. Using Lemma 1.60, it is not hard to see that this construction yields a Lie groupoid encoding the equivalence relation $x \sim y \Leftrightarrow \pi(x) = \pi(y)$. We mention two special cases of this construction:

- If $\pi = \mathrm{id}_M$, we obtain the *unit groupoid* $\mathfrak{u}(M) := (M \rightrightarrows M)$ (where all structure mappings are the identity). Clearly $\mathrm{Bis}(\mathfrak{u}(M)) = \{\mathrm{id}_M\}$.
- The map $\pi\colon M \to \{\bullet\}$ yields the *pair groupoid*, $\mathfrak{p}(M) = (M\times M \rightrightarrows M)$. We shall see in Exercise 6.2.2 that $\mathrm{Bis}(\mathfrak{p}(M)) \cong \mathrm{Diff}(M)$.

6.10 Example Consider a (left) Lie group action $\alpha\colon G\times M \to M$, $(g,m)\mapsto g\cdot m$. We form the *action groupoid* $\mathcal{A}_\alpha = (G\times M \rightrightarrows M)$, where $\mathbf{s}(g,m) := m$ and $\mathbf{t}(g,m) := \alpha(g,m)$. Now multiplication is defined as $(g,hm)\cdot(h,m) := (gh,m)$. The associated bisection group can be identified as

$$\mathrm{Bis}(\mathcal{A}_\alpha) = \{\sigma \in C^\infty(M,G) \mid m \mapsto \alpha(\sigma(m),m) \text{ is a diffeomorphism}\}. \quad (6.3)$$

Now $C^\infty(M,G) \to C^\infty(M,M)$, $f \mapsto \alpha_*(f\times \mathrm{id}_M)$ is smooth. Continuity of this map together with $\mathrm{Diff}(M) \subseteq C^\infty(M,M)$ yields $\mathrm{Bis}(\mathcal{A}_\alpha) \subseteq C^\infty(M,G)$. However, the bisections are not an open subgroup of the current group $C^\infty(M,G)$ (see §3.4) as the multiplication is $\sigma\star\tau(m) = \sigma(\tau(m)\cdot m)\cdot m$ instead of the pointwise product.

It is important to note that an action groupoid contains the same information as the group action and the quotient space. So instead of the ill-behaved quotient of the Lie group action

$$p\colon \mathrm{Imm}(\mathbb{S}^1,\mathbb{R}^2)\times\mathrm{Diff}(\mathbb{S}^1) \to \mathrm{Imm}(\mathbb{S}^1,\mathbb{R}^2)$$

from Example 6.4, one could instead work with the (infinite-dimensional) Lie groupoid $\mathcal{A}_p = (\mathrm{Imm}(\mathbb{S}^1,\mathbb{R}^2)\times\mathrm{Diff}(\mathbb{S}^1) \rightrightarrows \mathrm{Imm}(\mathbb{S}^1,\mathbb{R}^2))$ and carry out the geometric analysis on the groupoid instead of the quotient shape space. Since this quotient is of interest in shape analysis, Riemannian structures compatible with the Lie groupoid structure would then be needed to replace the metric on the quotient space. A suitable concept for such metrics has been worked out in del Hoyo and Fernandes (2018).

6.11 Remark Another interesting perspective on Lie groupoids is that they model symmetries which cannot be described by a global group action. A class of groupoids which fits well to this theme are the orbifold atlas groupoids (Moerdijk and Pronk, 1997). Recall that an *orbifold* is a manifold with mild singularities, that is, a Hausdorff space which is locally homeomorphic to a manifold

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

modulo a finite group of diffeomorphisms. The key point here is that the local group acting is allowed to change. As a visual example consider the sphere $\mathbb{S}^2$ where the upper half is rotated around the north pole by a rotation group of order p, while the lower half is rotated around the south pole by a rotation group of order q and the results are glued together. Topologically the space is still $\mathbb{S}^2$ but the manifold structure breaks down at the two fixed points. It has been shown that these structures are equivalent to certain Lie groupoids. We refrain from discussing the rather technical details and refer instead to Moerdijk and Pronk (1997) as well as Moerdijk and Mrčun (2003) for a detailed account.

The concept of a Lie groupoid carries over without any changes to infinite-dimensional settings (using submersions as defined in §1.7). For example, the action Lie groupoid modelling (6.4) is infinite dimensional. Let us mention further examples: In Beltiţă et al. (2019) Lie groupoids modelled on Banach spaces were studied. These arise naturally in studying certain pseudo-inverses in C^*-algebras. As a more concrete example of a genuine infinite-dimensional Lie groupoid consider the following.

6.12 Example Let $\mathcal{G} = (G \rightrightarrows M)$ be a finite-dimensional Lie groupoid and K a compact manifold. Then the pushforwards of the groupoid operations yield a Lie groupoid, called the *current groupoid* $C^\infty(K,\mathcal{G}) := (C^\infty(K,G) \rightrightarrows C^\infty(K,M))$. It is an easy exercise (Exercise 6.2.3) to verify that the current groupoid is a Lie groupoid. The theory for such groupoids was developed in Amiri et al. (2020). There it was shown that current groupoids inherit many structural properties from the finite-dimensional target groupoids.

However, it should be noted that from the rich theory available for finite-dimensional Lie groupoids (Mackenzie, 2005) virtually nothing is known for Lie groupoids modelled on general locally convex spaces. It is, for example, unclear as to whether the vertex groups always inherit a Lie group structure from the ambient groupoid.

6.13 Example Let (E,p,M,F) be a principal G-bundle, where E,M are Banach manifolds and G is a Banach Lie group (see Definition 3.52). Denote by $e \cdot g$ the right-G action on E and consider the diagonal G-action $(e,f)\cdot g := (e\cdot g, f\cdot g)$ on $E\times E$. Then the quotient $Q := (E\times E)/G$ is a manifold (with the unique structure turning the quotient map into a submersion). We obtain a Lie groupoid, called the *Gauge groupoid* $\mathrm{Gauge}(E) = (Q \rightrightarrows M)$, associated to the principal G-bundle. The groupoid is given by the following source, target and identity maps (see Exercise 6.2.5):

$$\mathbf{s}([e,f]) := p(f), \quad \mathbf{t}([e,f]) := p(e), \quad \mathbf{1}(m) := [u,u],$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

where $u \in p^{-1}(x)$ is arbitrary. Then we use the difference map $\delta\colon E\times_M E \to G$, $[e\cdot g, e] \mapsto g$ to define the multiplication

$$\mathbf{m}([e,f],[\tilde{e},\tilde{f}]) \coloneqq [e, f\delta(f,\tilde{e})], \text{ where } (f,\tilde{e}) \in E \times_M E.$$

So we can associate to every principal bundle (of Banach manifolds) a Lie groupoid. Conversely, one can show that the gauge groupoid uniquely identifies the principal bundle (see Mackenzie, 2005, Proposition 1.3.5). Hence a gauge groupoid contains the same information as a principal bundle. The bisection group $\mathrm{Bis}(\mathrm{Gauge}(E))$ is the automorphism group

$$\mathrm{Aut}(E,p,M) = \{f \in \mathrm{Diff}(E) \mid p\circ f \in \mathrm{Diff}(M), f(v\cdot g) = f(v)\cdot g, \text{ for all } v\in E, g\in G\}.$$

This group is known to be an infinite-dimensional Lie group (see Abbati et al., 1989) which contains the group of gauge transformations from Definition 3.56 as a (proper) Lie subgroup.

We have now seen in several examples that Lie groupoids can be used to formulate concepts from finite-dimensional differential geometry such as Lie group actions and principal bundles. Moreover, they come with an associated group, the bisection group, which in some instances can be identified with infinite-dimensional Lie groups. The next proposition shows that this is no accident.

6.14 Proposition *Assume that for a Lie groupoid $\mathcal{G}$, G is finite dimensional and M is compact. Then $\mathrm{Bis}(\mathcal{G})$ is a Lie group and $\mathbf{t}_*\colon \mathrm{Bis}(\mathcal{G}) \to \mathrm{Diff}(M)$, $\sigma \mapsto \mathbf{t}\circ\sigma$ is a Lie group morphism.*

Proof Recall from Example 3.5 that $\mathrm{Diff}(M)$ is an open submanifold of $C^\infty(M,M)$. Further, the pushforward $\mathbf{t}_*\colon C^\infty(M,G) \to C^\infty(M,M)$, $f \mapsto \mathbf{t}\circ f$ is smooth by Corollary 2.19. Since $\mathbf{s}\colon G \to M$ is a submersion, the Stacey–Roberts lemma, 2.24, asserts that $\mathbf{s}_*\colon C^\infty(M,G) \to C^\infty(M,M)$ is a submersion, whence the restriction $\theta \coloneqq \mathbf{s}_*|_{\mathbf{t}_*^{-1}(\mathrm{Diff}(M))}$ is a submersion. We deduce that $\theta^{-1}(\mathrm{id}_M) = \mathrm{Bis}(\mathcal{G})$ is a submanifold of $C^\infty(M,G)$. To see that this manifold structure turns $\mathrm{Bis}(\mathcal{G})$ into a Lie group, we rewrite the formulae (6.1) and (6.2) as follows:

$$\sigma \star \tau = \mathbf{m}_*(\mathrm{Comp}(\sigma, \mathbf{t}_*(\tau))), \tau) \qquad \sigma^{-1} = \mathbf{1}_* \circ \mathrm{Comp}(\sigma, \iota\circ \mathbf{t}_*(\sigma)),$$

where $\mathbf{m}$ is groupoid multiplication, $\mathbf{i}$ is groupoid inversion and ι the inversion in the Lie group $\mathrm{Diff}(M)$ (see Example 3.5). Since M is compact, pushforwards and the composition map are smooth by Proposition 2.23. In conclusion the group operations are smooth as composition of smooth mappings.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

As $\mathrm{Bis}(\mathcal{G}) \subseteq C^\infty(M,G)$ is a submanifold, the smoothness of $\mathbf{t}_*$ on $\mathrm{Bis}(\mathcal{G})$ follows from the smoothness of pushforwards on manifolds of mappings; Corollary 2.19. To see that $\mathbf{t}_*$ is a group morphism, we observe that

$$(\mathbf{t}_*(\sigma \star \tau)(x) = \mathbf{t}(\sigma(\mathbf{t}(\tau(x)))\tau(x)) = \mathbf{t}(\sigma(\mathbf{t}(\tau(x)))) = (\mathbf{t}_*(\sigma) \circ \mathbf{t}_*(\tau))(x). \quad \square$$

6.15 Remark The assumptions on $\mathcal{G}$ in the formulation of Proposition 6.14 are superfluous. The same proof (see Amiri and Schmeding, 2019, Proposition 1.3) works for any finite-dimensional Lie groupoid (dropping the compactness assumption on M), while in Schmeding and Wockel (2015, Theorem A) a proof for compact M but infinite-dimensional G was given (thus dropping the assumption on G). The latter proof is believed to generalise to non-compact M (and infinite-dimensional G).

6.16 Remark The Lie group structure of the bisections turns $\mathrm{Bis}(\mathfrak{p}(M)) \cong \mathrm{Diff}(M)$ into an isomorphism of Lie groups. Note however, that Proposition 6.14 cannot replace the classical construction of the Lie group structure on $\mathrm{Diff}(M)$ as we exploited this structure already in the proof of the proposition.

6.17 Remark The kernel of the Lie group morphism $\mathbf{t}_*\colon \mathrm{Bis}(\mathcal{G}) \to \mathrm{Diff}(M)$ is the *group of vertical bisections* $\mathrm{vBis}(\mathcal{G})$. Under certain assumptions on the Lie groupoid, it was shown in Schmeding (2020) that the vertical bisections form an infinite-dimensional Lie group.

If $\mathcal{G}$ is a gauge groupoid of some principal bundle, the vertical bisections coincide with the group of gauge transformations of the bundle. Moreover, in this case, the Lie group structure of the vertical bisections coincides with the Lie group structure on the group of gauge transformations; Remark 3.59.

Note that similar to the diffeomorphism group acting via evaluation on the underlying manifold, there is a canonical smooth action of the bisection group on the manifold of arrows of the groupoid.

6.18 Lemma *The evaluation map induces a Lie group action:*

$$\gamma\colon \mathrm{Bis}(\mathcal{G}) \times G \to G, \quad (\sigma, g) \mapsto \sigma(\mathbf{t}(g)) \cdot g.$$

Proof Setting in the definition, it is immediately clear that γ is a group action. Now rewrite γ as $\gamma(\sigma,g) = \mathbf{m}(\mathrm{ev}(\sigma, \mathbf{t}(g)), g), \quad \sigma \in \mathrm{Bis}(\mathcal{G}), g \in G$. Exploiting the smoothness of the evaluation map, Lemma 2.16, we see that the action is smooth. $\square$

With the help of the action one can identify the Lie algebra of the diffeomorphism group (see Exercise 6.2.6). We will focus here on global aspects of the theory and thus do not go into the details of the construction. However,

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

it should be remarked that the Lie algebra of the bisection group is closely connected to the infinitesimal level of Lie groupoid theory. To make sense of this, let us mention that every Lie groupoid admits an infinitesimal object called a Lie algebroid (its role is similar to that of a Lie algebra associated to a Lie group). A Lie algebroid is a vector bundle together with certain additional structures. Equivalently, a Lie algebroid can be formulated as a special type of Lie algebra, called Lie–Rinehart algebra. In the present case, the Lie–Rinehart algebra turns out to be the Lie algebra of the bisection group. This is left as Exercise 6.2.7 and we refer to Schmeding and Wockel (2015) as well as Mackenzie (2005) for more information.

Exercises

6.2.1 Let $\mathcal{G} = (G \rightrightarrows M)$ be a Lie groupoid. Show that:

(a) the domain of the multiplication $\mathbf{m}$ is a smooth manifold (whence it makes sense to require it to be smooth in the definition of a Lie groupoid);

(b) the unit map $\mathbf{1}\colon M \to G$ is a section of $\mathbf{s}$ and $\mathbf{t}$ and as a consequence, $\mathbf{1}$ is a smooth embedding (i.e. an immersion which is a homeomorphism onto its image);

(c) if only $\mathbf{s}$ is a submersion, so is $\mathbf{t}$ (vice versa if $\mathbf{t}$ is a submersion, so is $\mathbf{s}$). Hence the submersion requirements in the definition of a Lie groupoid can be weakened.

6.2.2 Let M be a manifold and $\alpha\colon G \times M \to M$ a Lie group action.

(a) Show that the bisection group of the pair groupoid $\mathfrak{p}(M)$ is isomorphic (as a group) to the diffeomorphism group $\mathrm{Diff}(M)$.

(b) Assume, in addition, that M is compact. Show that the group isomorphism from (a) becomes a Lie group isomorphism where the Lie group structure of $\mathrm{Bis}(\mathcal{G})$ is as in Proposition 6.14 and the one on $\mathrm{Diff}(M)$ as in Example 3.5.

(c) Work out the bisection group $\mathrm{Bis}(\mathcal{A}_\alpha)$ of the associated action groupoid. When is this group isomorphic to the current group $C^\infty(M,G)$?

6.2.3 Use the Stacey–Roberts Lemma, 2.24, to prove that the current groupoid $C^\infty(K,\mathcal{G})$ is a Lie groupoid for a finite-dimensional Lie groupoid $\mathcal{G}$.

6.2.4 Let $G \rightrightarrows M$ be a Lie groupoid such that G, M are manifolds modelled on Banach spaces. Show that the multiplication map $\mathbf{m}\colon G \times_M G \to G$ is a submersion. Deduce that the multiplication in every Banach Lie group is a submersion.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Hint: Since we are in the Banach setting, a submersion is a mapping which admits smooth local sections (see Exercise 1.7.5).

6.2.5 Let (E,p,M,F) be a principal G-bundle, where E,M are Banach manifolds and G is a Banach Lie group. We check that the associated gauge groupoid is a Lie groupoid. Show that:

(a) One can construct a manifold structure on the quotient $(E \times E)/G$ turning the quotient map $E \times E \to (E \times E)/G$ into a submersion.
Hint: Cover M by domains of sections of the submersion p and use the sections to construct charts for the manifold. For the submersion use Exercise 6.2.4.

(b) The structure maps $\mathbf{s},\mathbf{t},\mathbf{1},\mathbf{m}$ are smooth and that $\mathbf{s},\mathbf{t}$ are submersion. Conclude that the gauge groupoid is a Lie groupoid.
Hint: Assume p is a surjective submersion and q a smooth map between Banach manifolds. Then if $q \circ p$ is a submersion so is q (Margalef-Roig and Domínguez, 1992, Proposition 4.1.5).

6.2.6 Let $\mathcal{G} = (G \rightrightarrows M)$ be a Lie groupoid. Show that by applying the tangent functor T to every manifold and structure map of $\mathcal{G}$, one obtains a Lie groupoid $T\mathcal{G}$. One calls $T\mathcal{G}$ the *tangent (prolongation) groupoid* of $\mathcal{G}$.

6.2.7 In this exercise we identify the Lie algebra of the bisections $\mathrm{Bis}(\mathcal{G})$ as a Lie algebra of sections of a certain vector bundle. Note that this is precisely the algebra of sections induced by the Lie algebroid $\mathbf{L}(\mathcal{G})$ associated to $\mathcal{G}$. The Lie algebroid is the infinitesimal object associated to $\mathcal{G}$ (similar to the Lie algebra associated to a Lie group). As we have no need for a discussion of Lie algebroids, we will not discuss it but refer instead to Mackenzie (2005, 3.5).

(a) Exploit that $\mathbf{s}$ is a submersion, and use submersion charts to show that

$$T^{\mathbf{s}}G := \bigcup_{g\in G} T_g\,\mathbf{s}^{-1}(\mathbf{s}\,g) = \bigcup_{g\in G} \ker T_g\,\mathbf{s}$$

is a submanifold of TG and even a subvector bundle of TG.

(b) Use Exercise 1.7.3 to show that

$$\begin{aligned} T_{\mathbf{1}}\,\mathrm{Bis}(\mathcal{G}) &= \ker T_{\mathbf{1}}\,\mathbf{s}_* \\ &= \{f \in C^\infty(M,TG) \mid f(m) \in T_{\mathbf{1}(m)}G \text{ for all } m \in M\} \end{aligned}$$

and deduce that $T_{\mathbf{1}}\,\mathrm{Bis}(\mathcal{G}) \cong \Gamma(\mathbf{1}^*\,T^{\mathbf{s}}G)$ are vector spaces (where the right hand denotes sections of the pullback bundle $T^{\mathbf{s}}G$).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(c) Observe that $\mathbf{m}(g,\cdot)\colon \mathbf{s}^{-1}(\mathbf{s}\,g) \to \mathbf{s}(g)$, $h \mapsto gh$ is smooth for every $g \in G$ and show that every $X \in \Gamma(\mathbf{1}^* T^{\mathbf{s}}G)$ extends to a vector field on G via the formula

$$\overrightarrow{X}(g) \coloneqq T(\mathbf{m}(g,\cdot))(X(\mathbf{t}(g))).$$

Prove that $X = \overrightarrow{X} \circ \mathbf{1}$, whence the linear map $\Gamma(\mathbf{1}^* T^{\mathbf{s}}G) \to \mathcal{V}(G)$, $X \mapsto \overrightarrow{X}$ must be injective and we can define a Lie bracket on $\Gamma(\mathbf{1}^* T^{\mathbf{s}}G)$ via $[X,Y] \coloneqq [\overrightarrow{X},\overrightarrow{Y}] \circ \mathbf{1}$ (where the Lie bracket on the right is the Lie bracket of vector fields).

(d) Adapt the proof identifying the Lie bracket of the algebra for the diffeomorphism group to bisection groups, that is, show that if X^R is a right-invariant vector field on $\mathrm{Bis}(\mathcal{G})$ then the vector field $X^R \times 0_G \in \mathcal{V}(\mathrm{Bis}(\mathcal{G}) \times G)$ is γ-related to $\overrightarrow{X(\mathbf{1})}$. Deduce from this that the Lie bracket can be identified with the negative of the bracket from (c).

6.2.8 Let $\mathcal{G}_1, \mathcal{G}_2$ be Lie groupoids. A *morphism of Lie groupoids* is a pair of smooth maps $F\colon G_1 \to G_2$ and $f\colon M_1 \to M_2$ such that $\mathbf{s}_2 \circ F = f \circ \mathbf{s}_1$, $\mathbf{t}_2 \circ F = f \circ \mathbf{t}_1$ and $F(gh) = F(g)F(h)$ (whenever, $g, h \in G_1$ are composable). If $f = \mathrm{id}_{M_1}$, we say F is a *morphism over the identity*. Show that every morphism F over the identity induces a Lie group morphism $F_*\colon \mathrm{Bis}(\mathcal{G}_1) \to \mathrm{Bis}(\mathcal{G}_2)$, $\sigma \mapsto F \circ \sigma$.

Remark: So far we have avoided Lie groupoid morphisms as Lie groupoids and general morphisms exhibit a more complicated interplay as they form a 2-category (thus the moniker higher geometry). We will not delve into the details of this construction.

6.3 (Re-)construction of a Lie Groupoid from Its Bisections

We shall now consider whether a Lie groupoid is determined by its group of bisections or can even be reconstructed from this group. For the reconstruction, we consider again the Lie group action of the bisections on the manifold of arrows. This action turns out to be a submersion.

6.19 Proposition *Let $\mathcal{G}$ be a finite-dimensional Lie groupoid. Then the following mappings are submersions:*

(a) $\gamma_m\colon \mathrm{Bis}(\mathcal{G}) \to \mathbf{s}^{-1}(m)$, $\sigma \mapsto \sigma(m)$, for all $m \in M$,
(b) $\mathrm{ev}\colon \mathrm{Bis}(\mathcal{G}) \times M \to G$, $\sigma \mapsto \sigma(m)$,
(c) $\gamma\colon \mathrm{Bis}(\mathcal{G}) \times G \to G$, $(\sigma, g) \mapsto \sigma(\mathbf{t}(g)) \cdot g$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Proof (a) In Exercise 2.3.4 we saw that the tangent of $\mathrm{ev}_m\colon C^\infty(M,G) \to G$, $f \mapsto f(m)$ is given by $T_f \mathrm{ev}_m\colon C^\infty_f(M,TG) \to T_{f(m)}G$, $h \mapsto h(m)$. By assumption, $\mathbf{s}^{-1}(m) \subseteq G$ is a finite-dimensional manifold, whence 1.56 shows that it suffices to prove that for each $\sigma \in \mathrm{Bis}(\mathcal{G})$ the tangent map $T_\sigma \mathrm{Bis}(\mathcal{G}) \to T_{\sigma(m)}\mathbf{s}^{-1}(m)$ is surjective.[1] By construction $(\mathbf{s}_*)^{-1}(\mathrm{id}_M) \cap (\mathbf{t}_*)^{-1}(\mathrm{Diff}(M)) = \mathrm{Bis}(\mathcal{G}) \subseteq C^\infty(M,G)$ and since $\mathbf{s}_*$ is a submersion we have, with Exercise 1.7.3 and arguments as in Exercise 6.2.6(b), that

$$T_\sigma \mathrm{Bis}(\mathcal{G}) = \ker T_\sigma(\mathbf{s}_*) = \ker(T\,\mathbf{s})_*|_{T_\sigma C^\infty(M,G)} \cong \Gamma(\sigma^* T^{\mathbf{s}}G) \subseteq C^\infty_\sigma(M,G).$$

This shows that $T_\sigma \gamma_m = T_\sigma \mathrm{ev}_m|_{T_\sigma \mathrm{Bis}_\mathcal{G}}$ is surjective as every element of $T^{\mathbf{s}}_{\sigma(m)}G$ can be written as $X(\sigma(m))$ for some $X \in \Gamma(T^{\mathbf{s}}G) \cong \Gamma(\sigma^* T^{\mathbf{s}}G)$ (see Exercise 6.3.2). We deduce that γ_m is a submersion.

(b) The proof turns out to be quite involved and involves a reduction step to the case already dealt with in (a). Note that this is not obvious as $T\,\mathbf{s}^{-1}(m)$ is, in general, properly contained in TG. For similar reasons we cannot deduce the submersion property of ev from the submersion property of $\mathrm{ev}\colon C^\infty(M,G) \times M \to G$. For the proof and the necessary details we refer to Schmeding and Wockel (2016, Proposition 2.8 and Corollary 2.10).

(c) First note that we can write $\gamma(\sigma,g) = \mathbf{m}(\mathrm{ev}(\sigma,\mathbf{t}(g)),g)$ and both $\mathbf{t}$ and $\mathbf{m}$ are submersions (see Exercise 6.2.4). Hence γ is a submersion as $\mathrm{ev}\colon \mathrm{Bis}(\mathcal{G}) \times M \to G$ is a submersion by (b). □

We now have a submersion $\mathrm{ev}\colon \mathrm{Bis}(\mathcal{G}) \times M \to G$ which is a Lie group action of the bisections on the arrow manifold of the groupoid from which we constructed the bisections. Note that the unit embeds M as a submanifold of G and $\mathbf{t}_*(\mathrm{Bis}(\mathcal{G})) \subseteq \mathrm{Diff}(M)$ Exercise 6.2.1 (b). This observation allows us to prolong the action of the bisections to an action on M:

$$A\colon \mathrm{Bis}(\mathcal{G}) \times M \to M, \quad (\sigma,m) \mapsto \mathbf{t}(\sigma(m)).$$

We will now show that the action groupoid constructed from this action determines (under certain conditions) the Lie groupoid $\mathcal{G}$.

6.20 Definition Let $\mathcal{G} = (G \rightrightarrows M)$ be a Lie groupoid and $A\colon \mathrm{Bis}(\mathcal{G}) \times M \to M$ the canonical Lie group action of the bisections on the units. Then we call the action groupoid $\mathrm{Bis}(\mathcal{G}) \ltimes M := (\mathrm{Bis}(\mathcal{G}) \times M \rightrightarrows M)$ the *bisection action groupoid*. Furthermore, the map $\mathrm{ev}\colon \mathrm{Bis}(\mathcal{G}) \times M \to G$ induces a Lie groupoid morphism **ev** over the identity.[2]

[1] While the above formula immediately shows that ev_m is a submersion, we cannot directly conclude this for γ_m without identifying the subspace of $C^\infty_\sigma(M,G)$ associated to $T_\sigma \mathrm{Bis}(\mathcal{G})$.

[2] A Lie groupoid morphism over the identity is a smooth map $f\colon G \to G'$ (for groupoids

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

The question is now of course whether the bisection action groupoid completely determines the Lie groupoid from which the bisections were constructed. In general, this will not be the case as there will not be enough bisections to obtain all elements in the arrow manifold of a Lie groupoid.

6.21 Example Let M be a compact manifold with two connected components $M = M_1 \sqcup M_2$ such that M_1 and M_2 are *not* diffeomorphic. We have seen in the previous chapter that for the pair groupoid $\mathfrak{p}(M) = (M \times M \rightrightarrows M)$, the bisection group can be identified as $\mathrm{Bis}(\mathfrak{p}(M)) \cong \mathrm{Diff}(M)$. Now consider an element $(m_1, m_2) \in M$ such that $m_1 \in M_1$ and $m_2 \in M_2$. If there were a bisection σ such that $\sigma(m_2) = (m_1, m_2)$, this would imply that there must be a diffeomorphism $\phi\colon M \to M$ such that $\phi(m_2) = m_1$. As this entails $\phi(M_2) = M_1$ (since diffeomorphisms permute the connected components of a manifold), this is clearly impossible. We conclude that for every pair such that the elements come from different components, there cannot be a bisection through this element of the Lie groupoid.

So if there should be any hope that the bisection group identifies the Lie groupoid from which it was derived, we need to require that there are enough bisections in the following sense.

6.22 Definition A Lie groupoid $\mathcal{G} = (G \rightrightarrows M)$ is said to have *enough bisections* if for every $g \in G$ there exists a bisection $\sigma_g \in \mathrm{Bis}(\mathcal{G})$ with $\sigma_g(\mathbf{s}(g)) = g$.

Fortunately, sufficient conditions for a Lie groupoid to possess enough bisections are known. Indeed it turns out that the deficiency pointed out in Example 6.21 is caused by a lack of connectedness. This can be remedied by requiring that the groupoid $\mathcal{G}$ is *source connected*, that is, for every $m \in M$ the source fibre $\mathbf{s}^{-1}(m)$ is connected.

6.23 Remark If $\mathcal{G}$ is the pair groupoid of a manifold, $\mathrm{Bis}(\mathcal{G}) \cong \mathrm{Diff}(M)$, the groupoid is source connected if and only if M is connected. Our next result will entail that $\mathrm{Diff}(M)$ acts transitively on M. We remark that connectedness of M was required in the statement of the Takens–Filipkiewicz result in §6.1. Note that transitivity of the group action was an essential ingredient in the proof of the result.

6.24 Lemma *If $\mathcal{G}$ is source connected, then $\mathcal{G}$ has enough bisections.*

$G \rightrightarrows M$ and $G' \rightrightarrows M$ which relates the structural maps of the groupoids, that is, $\mathbf{s}' \circ f = \mathbf{s}$, $\mathbf{t} \circ f = \mathbf{t}$, $\mathbf{m}' \circ (f, f) = f \circ \mathbf{m}$ and $\mathbf{i}' \circ f = f \circ \mathbf{i}$. We verify the conditions for ev in Exercise 6.3.2.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Proof The image $\mathcal{U} := \mathrm{ev}(\mathrm{Bis}(\mathcal{G}) \times M)$ contains the image of the object inclusion $\mathbf{1}$, that is, $\mathbf{1}(m) \in \mathcal{U}$ for all $m \in M$. Define for $m \in M$ the set $\mathcal{U}_m = \mathcal{U} \cap \mathbf{s}^{-1}(m)$ and note that $\mathcal{U}_m = \mathrm{ev}(\mathrm{Bis}(\mathcal{G}) \times \{m\})$. Clearly the set $\mathcal{U}$ contains $\mathbf{1}(M)$ and forms a subgroupoid $\mathcal{U} \rightrightarrows M$ of $\mathcal{G}$ of $G \rightrightarrows M$ (see Exercise 6.3.3). As submersions are open maps, $\mathcal{U}$ is an open Lie subgroupoid of $\mathcal{G}$. Now as $\mathrm{ev}_m \colon \mathrm{Bis}(\mathcal{G}) \to \mathbf{s}^{-1}(m)$ is a submersion by Proposition 6.19(a) we infer that $\mathcal{U}_m$ is an open subset of $\mathbf{s}^{-1}(m)$. However, $\mathcal{U}_m$ is also closed: The complement $\mathbf{s}^{-1}(m) \setminus \mathcal{U}_m$ is the union $\bigcup_{g \in \mathbf{s}^{-1}(m) \setminus \mathcal{U}_m} \mathcal{U}_{\mathbf{t}(g)} \cdot g$. As $\mathcal{U}_{\mathbf{t}(m)} \trianglelefteqslant \mathbf{s}^{-1}(\mathbf{t}(g))$, we see that $\mathcal{U}_{\mathbf{t}(g)} \cdot g$ is open, whence $\mathcal{U}_m$ is also closed. We deduce that the clopen set $\mathcal{U}_m \subseteq \mathbf{s}^{-1}(m)$ equals $\mathbf{s}^{-1}(m)$ as $\mathcal{G}$ is source connected. □

If $\mathcal{G}$ has enough bisections, the evaluation map $\mathrm{ev} \colon \mathrm{Bis}(\mathcal{G}) \times M \to G$ becomes a surjective submersion. Hence the Lie groupoid structure of $\mathcal{G}$ is completely determined by the group of bisections. One can moreover show that the original groupoid is a quotient groupoid of the bisection action groupoid in this case (see Schmeding and Wockel, 2016, Theorem B). As we do not wish to introduce deeper concepts in groupoid theory, we do not go into details concerning this result. The main upshot, however, is that for a Lie groupoid with enough bisections, the groupoid is uniquely determined by the action of the bisection group.

6.25 Remark The results presented so far in this section are reconstruction results. This means that starting from a Lie groupoid, we can recover the Lie groupoid (under certain topological assumptions) as the quotient of a Lie groupoid we cook up from the action of the bisection group. One can of course ask whether there are construction results for Lie groupoids which do not require starting from a Lie groupoid. Instead, one would like to start from an action of a suitable infinite-dimensional group and construct a Lie groupoid such that, in the case we started with the canonical action of a bisection group, one would recover the Lie groupoid. At least partial answers to these questions exist. We refer to Schmeding and Wockel (2016), where transitive pairs, that is, a principal bundle version of Klein geometries (Sharpe, 1997, Chapter 3), are proposed as a starting point for a construction result.

Exercises

6.3.1 Let $\mathcal{G} = (G \rightrightarrows M)$ be a Lie groupoid. Show that the image of the canonical action $E := \mathrm{ev}(\mathrm{Bis}(\mathcal{G}) \times M) \subseteq G$ is closed under multiplication and inversion in G. Hence with the induced structure maps we obtain a subgroupoid $E \rightrightarrows M$ of $\mathcal{G}$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

6.3.2 Prove that $\mathrm{ev}\colon \mathrm{Bis}(\mathcal{G}) \times M \to G$ induces a Lie groupoid morphism $\mathrm{Bis}(\mathcal{G}) \ltimes M \to G$ over the identity.

6.3.3 Let $\pi\colon E \to M$ be a finite-dimensional vector bundle. Show that for every $e \in E$ there is $X^e \in \Gamma(E)$ with $X^e(\pi(e)) = e$.
Hint: Construct locally in trivialisations using bump functions. Note that the assumption of being finite dimensional can be replaced by requiring the existence of suitable bump functions.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

7

Euler–Arnold Theory: PDEs via Geometry

7.1 Introduction

In this chapter we shall give an introduction to the Euler–Arnold theory for partial differential equations (PDEs). The main idea of this theory is to reinterpret certain PDEs as smooth ordinary differential equations (ODEs) on infinite-dimensional manifolds. One advantage of this idea is that the usual solution theory for ODEs can be used to establish properties for the PDE under consideration. This principle has been successfully applied to a variety of PDEs arising, for example, in hydrodynamics. Among these are the Euler equations for an ideal fluid, the Camassa–Holm equation, the Hunter–Saxton equation and the inviscid Burgers equation. We refer to Khesin and Wendt (2009, p. 34) for a much longer list of physically relevant PDEs which fit into this setting.

As in the rest of the book, we shall only work with smooth functions. This is rather unnatural for solutions of partial differential equations but allows us to avoid spaces and manifolds of finitely often differentiable mappings. From the theoretical point of view, this is problematic (at least considering the results we are after) and we will comment on the 'correct setting' at the end of this chapter. Before we begin, recall the relation between the energy of a curve and it being a geodesic (see §4.2).

7.1 Definition Let M be a manifold and for $x, y \in M$ we define the closed submanifold

$$C^\infty_{x,y}([0,1],M) := \{c \in C^\infty([0,1],M) \mid c(0) = x, c(1) = y\} \subseteq C^\infty([0,1],M)$$

of curves from x to y (see Exercise 7.1.1). A smooth map $p\colon]-\delta,\delta[\times[0,1] \to M$ is a *smooth variation* of $c \in C^\infty_{x,y}([0,1],M)$ if $p(0,t) = c(t)$, for all $t \in [0,1]$ and $p(s,\cdot) \in C^\infty_{x,y}([0,1],M)$, for all $s \in]-\delta,\delta[$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

With the notation in place, we can now formulate the following standard result from Riemannian geometry for weak Riemannian metrics which admit a metric spray.

7.2 Proposition *Let (M,g) be a (weak) Riemannian manifold which admits a metric spray S. Then $c\colon [0,1] \to M$ is a geodesic if and only if it extremises the energy*

$$\mathrm{En}(c) = \frac{1}{2}\int_a^b g_{c(t)}(\dot{c}(t),\dot{c}(t))\mathrm{d}t,$$

that is, if for each smooth variation $p\colon]-\delta,\delta[\times[0,1] \to M$ of c we have

$$\left.\frac{d}{ds}\right|_{s=0} \mathrm{En}(p(s,\cdot)) = 0.$$

Proof Pick a smooth variation q of a curve c with $h(t) \coloneqq \left.\frac{\partial}{\partial s}\right|_{s=0} q(s,t)$. We compute the derivative by exploiting the formula for the derivative of the energy from Lemma C.17 in a local chart (U,φ). Suppressing again the identifications in the notation we find for $\left.\frac{d}{ds}\right|_{s=0}\mathrm{En}(q(s))$ the formula

$$\begin{aligned}
&\int_0^1 \frac{1}{2}d_1 g_U(c,c'(t),c'(t);h) - d_1 g_U(c(t),h(t),c'(t);c'(t)) \\
&\qquad - g_U(c(t),h(t),c''(t))\mathrm{d}t \\
&\overset{(4.6)}{=} \int_0^1 g_U(c(t),B_U(c(t),c'(t)),c'(t),h(t)) - g_U(c(t),h(t),c''(t))\mathrm{d}t \\
&= \int_0^1 g_U(c(t),B_U(c(t),c'(t),c'(t)) - c''(t),h(t))\mathrm{d}t \\
&\overset{(4.13)}{=} \int_0^1 -g_U(h(t),\nabla_{\dot{c}(t)}\dot{c}(t),h(t))\mathrm{d}t.
\end{aligned}$$

Hence $g_U(h(t),\nabla_{\dot{c}(t)}\dot{c}(t),h(t))$ needs to vanish for every h. Since the only element which gets annulled by all $g(h,\cdot)$ is 0, we conclude that $\left.\frac{d}{ds}\right|_{s=0}$ $\mathrm{En}(p(s,\cdot)) = 0$ holds for all smooth variations if and only if $\nabla_{\dot{c}}\dot{c} = 0$, that is, c is a geodesic. □

7.3 Remark Proposition 7.2 shows that geodesics are critical points of the energy. In §4.3 we have taken the perspective that geodesics are locally length minimising, that is, critical points of the length. However, the energy depends on the parametrisation while the length does not. To reconcile this, note that the Cauchy–Schwarz inequality yields equality of energy and length if $g_c(\dot{c},\dot{c})$ is constant. In Exercise 7.1.2 we will see that every geodesic satisfies this property, whence our definition of geodesic comes with a preferred parametrisation which makes both points of view equivalent.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

The variational approach will enable us to identify the geodesic equations of infinite-dimensional Riemannian metrics. Before we turn to these results, let us fix some notation for this and later sections.

Concerning Partial Derivatives Let $p\colon]-\delta,\delta[\times[0,1] \to M$ be a smooth variation. Assume that ∇ is the metric derivative of the Riemannian manifold (M,g). For $s \in]-\delta,\delta[$ and $t \in [0,1]$ we denote by ∂s and ∂t the constant unit vector field (on $]-\delta,\delta[$ resp. on $[0,1]$). For p, we denote by $\frac{\partial}{\partial s}p(s,t) \coloneqq Tp(s,t)(1,0)$ and $\frac{\partial}{\partial t}p(s,t) \coloneqq Tp(s,t)(0,1)$ the partial derivatives as vector fields along p. Hence we can consider, for example, $\frac{\nabla}{\partial s}\frac{\partial}{\partial t}p$ as a vector field along p (in the sense of Definition 5.4). Moreover, we note that as ∇ is the metric derivative (hence torsion free), Klingenberg (1995, Proposition 1.5.8.i) implies that

$$\frac{\nabla}{\partial s}\frac{\partial}{\partial t}p = \frac{\nabla}{\partial t}\frac{\partial}{\partial s}p. \tag{7.1}$$

7.4 Example (The inviscid Burgers equation) Recall from Corollary 2.8 that the diffeomorphism group $\mathrm{Diff}(\mathbb{S}^1) \subseteq C^\infty(\mathbb{S}^1,\mathbb{S}^1)$ is an open submanifold. Fix some Riemannian metric g on $\mathbb{S}^1$. We exploit that $\mathbb{S}^1 \subseteq \mathbb{R}^2$ is an embedded submanifold and apply Remark 5.9(c) to deduce from Proposition 5.8 that the L^2-metric on $C^\infty(\mathbb{S}^1,\mathbb{S}^1)$ admits a metric spray. The same then holds for its restriction to $\mathrm{Diff}(\mathbb{S}^1)$:

$$g^{L^2}(X,Y) = \int_{\mathbb{S}^1} g_{\varphi(\theta)}(X(\theta),Y(\theta))\mathrm{d}\theta,$$

where we exploited the identification $T_\varphi \mathrm{Diff}(\mathbb{S}^1) \cong \{X\circ\varphi \mid X \in \mathcal{V}(\mathbb{S}^1)\}$. Now pick a smooth variation $c\colon]-\delta,\delta[\times[0,1] \to \mathrm{Diff}(\mathbb{S}^1)$ of some curve (which we also denote by c). Set $v(t) \coloneqq \frac{\partial}{\partial s}\big|_{s=0} c(s,t)$ and note that taking the derivative with respect to s coincides with taking the derivative with respect to the unit vector field ∂s. We can now compute the variation of the energy

$$\begin{aligned}
\frac{d}{ds}\Big|_{s=0}\mathrm{En}(c(s,\cdot)) &= \frac{1}{2}\int_0^1\int_{\mathbb{S}^1}\frac{\partial}{\partial s}\Big|_{s=0} g\left(\frac{\partial}{\partial t}c(s,t)(x),\frac{\partial}{\partial t}c(s,t)(x)\right)\mathrm{d}\theta\mathrm{d}t\\
&\overset{(5.3)}{=} \int_0^1\int_{\mathbb{S}^1} g\left(\frac{\nabla}{\partial s}\frac{\partial}{\partial t}c(s,t)(x),\frac{\partial}{\partial t}c(s,t)(x)\right)\Big|_{s=0}\mathrm{d}\theta\mathrm{d}t\\
&\overset{(7.1)}{=} \int_0^1 g^{L^2}\left(\frac{\nabla}{\partial t}\frac{\partial}{\partial s}c(s,t)(x),\frac{\partial}{\partial t}c(s,t)(x)\right)\Big|_{s=0}\mathrm{d}t\\
&= -\int_0^1 g^{L^2}\left(v(t),\frac{\nabla}{\partial t}\frac{\partial}{\partial t}c(t)\right)\mathrm{d}t,
\end{aligned}$$

where the last equality is due to a usual integration by parts argument. In particular, we see that if $c\colon [0,1] \to \mathrm{Diff}(\mathbb{S}^1)$ should be a geodesic, it must satisfy the pointwise equation

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$\frac{\nabla}{\partial t}\frac{\partial}{\partial t}c^\wedge(t,\theta) = 0, \quad \text{for all } \theta \in \mathbb{S}^1. \tag{7.2}$$

To connect this equation to objects on the finite-dimensional manifold $\mathbb{S}^1$, we construct a (time-dependent) vector field on $\mathbb{S}^1$ from the curve of diffeomorphisms $c(t)$ by setting

$$u(t,\theta) = \frac{\partial}{\partial t}c^\wedge(t, c^{-1}(t,\theta)), \tag{7.3}$$

where the inverse is the inverse in $\mathrm{Diff}(M)$. In other words, $c(t)(\theta)$ is the flow of the time-dependent vector field u, that is, $\frac{\partial}{\partial t}c(t)(\theta) = u(t, c(t)(\theta))$. Plugging this definition into the left-hand side of (7.2), the chain rule and a quick computation yield the following statement.

7.5 Lemma *Let $c\colon [0,1] \to \mathrm{Diff}(\mathbb{S}^1)$ be smooth and the flow of the time-dependent vector field u on $\mathbb{S}^1$, cf. (7.3). Then*

$$\frac{\nabla}{\partial t}\frac{\partial}{\partial t}c^\wedge(t,\theta) = \frac{\nabla}{\partial t}(u(t,c^\wedge(t,\theta))) = \frac{\partial u}{\partial t}(t,c^\wedge(t,\theta)) + \nabla_u u(t,c^\wedge(t,\theta)), \tag{7.4}$$

where $\nabla_u u$ denotes the covariant derivative of u against itself for fixed t and we interpret $\frac{\partial u}{\partial t}(t,\theta)$ as a partial derivative of $u(\cdot,\theta)$ in $T_\theta\mathbb{S}^1$ for every fixed θ.

Note that (7.4) is central to the idea of the Euler–Arnold theory (whence we promoted it to its own lemma) and holds in similar form if one replaces $\mathbb{S}^1$ by an arbitrary smooth compact manifold M. To distinguish the interpretation of $\frac{\partial u}{\partial t}(t,\theta)$ from the usual partial derivative of a smooth variation, let us write $\partial_t u$ for this derivative. We conclude from Lemma 7.5 and (7.2) that c is a geodesic of the L^2-metric if and only if the associated vector field u solves the inviscid Burgers (or Hopf) equation

$$\partial_t u + \nabla_u u = 0. \tag{7.5}$$

Burgers' equation is a partial differential equation ($\nabla_u u$ takes derivatives of u) which is miraculously equivalent to an ordinary differential equation (the geodesic equation) on the infinite-dimensional group $\mathrm{Diff}(\mathbb{S}^1)$. Note that (7.5) is connected to the classical Burgers equation $u_t + 3uu_x = 0$ (subscripts denoting partial derivatives) from Example 4.15. This will be made explicit in Exercise 7.1.3: Since $\mathbb{S}^1 \subseteq \mathbb{R}^2$ is an embedded submanifold, we endow $\mathbb{S}^1$ with the pullback metric induced by the Euclidean inner product $g_x(v,w) \coloneqq \langle v, w\rangle$ (by identifying $T_x\mathbb{S}^1 \subseteq \mathbb{R}^2$). Working out the covariant derivatives, a canonical identification shows that the Burgers equation (7.4) coincides with the classical Burgers equation $u_t + uu_x = 0$ (which is up to scaling equivalent to Example 7.4).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

7.6 Remark The derivation of Burgers' equation as a geodesic equation on $\mathbb{S}^1$ did not exploit any special structure of $\mathbb{S}^1$. It just enabled us to make sense of the integrals without recourse to integration against volume forms. Thus the same argument carries over without any change to an arbitrary compact manifold M. There the inviscid Burgers equation

$$\partial_t u + \nabla_u u = 0$$

makes sense (with respect to the covariant derivative induced by the metric) and is the geodesic equation of the L^2-metric (cf. (7.8) below) on $\mathrm{Diff}(M)$.

In the next section we will systematically investigate the mechanism to associate a geodesic equation to certain partial differential equations.

Exercises

7.1.1 Let (M,g) be a (weak) Riemannian manifold (with metric derivative ∇) and denote by $C^\infty([0,1],M)$ the space of smooth curves with the manifold structure from Appendix C.4. Fix $x,y \in M$. Show that:

(a) $C^\infty_{x,y}([0,1],M) = \{c \in C^\infty([0,1],M) \mid c(0) = x, c(1) = y\}$ is a closed submanifold of $C^\infty([0,1],M)$.
Hint: Consider a canonical chart for the manifold of mappings. Show that the model space splits for every $c \in C^\infty_{x,y}([0,1],M)$.

(b) $p\colon]-\delta,\delta[\times[0,1] \to M$ is a smooth variation of $c \in C^\infty_{x,y}([0,1], M)$ if and only if $p^\vee\colon]-\delta,\delta[\to C^\infty_{x,y}([0,1],M)$ is smooth.

(c) The energy En restricts to a smooth function on $C^\infty_{x,y}([0,1],M)$ and prove the following analogue of Proposition 7.2: A curve c is a geodesic connecting x and y if and only if $d\mathrm{En}(c;\cdot)$ vanishes on $C^\infty_{x,y}([0,1],M)$.

7.1.2 Let (M,g) be a weak Riemannian manifold which allows a metric spray S and an associated metric derivative ∇. Show that a geodesic $c\colon [0,1] \to M$ is a curve of constant speed, that is, $g_c(\dot c,\dot c)$ is constant.
Hint: Use 4.29 to show that $\frac{d}{dt} g_c(\dot c,\dot c)$ vanishes.

7.1.3 Show that in Example 7.4 we can rewrite (7.5) in traditional notation as

$$u_t + uu_x = 0, \quad \text{where } u\colon]-\delta,\delta[\times[0,2\pi] \to \mathbb{R}.$$

In addition, show that in this setting the flow η of u then satisfies $\frac{\partial^2}{\partial t^2}\eta = 0$.
Hint: Use the idea that the tangent bundle of $\mathbb{S}^1$ is trivial together with Example 4.33.

7.1.4 Prove Lemma 7.5.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

7.2 The Euler Equation for an Ideal Fluid

We shall exhibit the general principle first for the classical example considered by Arnold (1966): the Euler equation for an inviscid incompressible fluid on a Riemannian manifold. It describes the development of a fluid occupying the manifold M under certain assumptions. Let us first fix some notation.

Conventions In this section we will denote by (M, g) a compact (thus finite-dimensional) orientable Riemannian manifold.

- For a (time-dependent) vector field u, we write $\partial_t u(t,x)$ for the partial t-derivative in $T_x M$ (thus not taking values in $T(TM)$!).
- Since M is orientable, it admits a volume form μ induced by g. Further, we denote by $\operatorname{div} X$ the *divergence* of a vector field $X \in \mathcal{V}(M)$ and by $\operatorname{grad} f$ the *gradient* of a smooth function $f\colon M \to \mathbb{R}$ (see Appendix E.3).

We will not derive Euler's equations here from first principles, but refer to Modin (2019) for an account together with a history of the problem.

7.7 (Euler equation) The *Euler equation* for an incompressible fluid is

$$\begin{cases} \partial_t u(t,m) + \nabla_u u(t,m) = -\operatorname{grad} p, & \\ \operatorname{div} u(t,\cdot) = 0 & \text{for all } t, \\ u(0,\cdot) = u_0 & \text{with } \operatorname{div} u_0 = 0, \end{cases} \tag{7.6}$$

where the function $p\colon \mathbb{R}\times M \to \mathbb{R}$ is interpreted as 'pressure'. Euler's equation searches for a (time-dependent) vector field u on M. The condition, $\operatorname{div} u(t,\cdot) = 0$, that is, that u is divergence-free, is the condition enforcing the incompressibility of the fluid.

In (7.6) we seek a vector field, whence one says that the equation is in *Eulerian form*.

7.8 Remark Apart from the incompressibility condition, Euler's equation is similar to Burgers' equation. Indeed the only difference on the PDE level is the right-hand side which is given by the gradient of a pressure function. We will see later that the gradient acts as a Lagrange multiplier enforcing the incompressibility condition (in general, the term $\nabla_u u$ will not be divergence-free).

Again Euler's equation, just like Burgers' equation, is formulated on a finite-dimensional manifold and has a priori nothing to do with infinite-dimensional geometry. However, we will change the perspective to uncover the connection to infinite-dimensional geometry. The idea is similar to what we did for Burgers' equation (but we will now start with the vector field rather than a flow).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

From the Eulerian to the Lagrangian Perspective

Let us consider a time-dependent smooth vector field $u\colon I \times M \to TM$ on a compact interval I. Then recall (e.g. from Lang, 1999, §IV.1, where we exploit M being compact) that the flow for u is a mapping $\eta\colon I \times M \to M$ such that

$$u(t,\eta(t,m)) = \frac{\partial}{\partial t}\eta(t,m), \text{ for all } t \in I,\ m \in M. \tag{7.7}$$

Furthermore, it is well known that for each $t \in I$ the flow $u(t,\cdot)$ is a diffeomorphism of M. The equation (7.7) or the equivalent equation (7.3) is sometimes called the *reconstruction equation*. Observe now that instead of constructing a vector field u which solves the Euler equation (7.6), we can construct its flow. Searching for the flow whose associated vector field solves the PDE is called the *Lagrangian perspective* on the PDE. If u is now divergence-free, $\operatorname{div} u = 0$, this implies that $\eta_*\mu = \mu$, that is, for every t, the diffeomorphism $\eta(t,\cdot)$ leaves the volume form invariant. As η is smooth, the exponential law, Theorem 2.12, allows us to reinterpret the flow $\eta\colon I \times M \to M$ as a smooth curve $\eta^\vee\colon I \to \operatorname{Diff}(M) \subseteq C^\infty(M,M)$. Now the incompressibility condition shows that $\eta^\vee$ takes its values in the closed Lie subgroup $\operatorname{Diff}_\mu(M) \subseteq \operatorname{Diff}(M)$ of volume-preserving diffeomorphisms.

Our aim is again to connect the finite-dimensional PDE to the infinite-dimensional Riemannian geometry induced by the L^2-metric. Let us briefly recall its definition.

7.9 Definition (L^2-metric on the diffeomorphism group) Let (M,g) be a compact Riemannian manifold[1]. We define a weak Riemannian metric on $\operatorname{Diff}(M)$ via

$$g^{L^2}_\varphi(X\circ\varphi, Y\circ\varphi) = \int_M g_{\varphi(m)}(X(\varphi(m)), Y(\varphi(m)))\mathrm{d}\mu(m). \tag{7.8}$$

Here $X,Y \in \mathcal{V}(M), \varphi \in \operatorname{Diff}(M)$ and we exploited $\operatorname{Diff}(M)$ being a Lie group, whence its tangent bundle is trivial, that is, $T\operatorname{Diff}(M) \cong \operatorname{Diff}(M) \times \mathcal{V}(M)$ (where the diffeomorphism is induced by right translation). It follows directly from the rules of integration that g^{L^2} is a right-invariant Riemannian metric (see 4.13) on the subgroup $\operatorname{Diff}_\mu(M)$ of volume-preserving diffeomorphisms (but not on $\operatorname{Diff}(M)$!). Moreover, in Exercise 7.3.2 we will see that the weak Riemannian metric admits a metric spray and a covariant derivative. With more work, one can also establish this for the restriction of the L^2-metric to the closed Lie subgroup $\operatorname{Diff}_\mu(M)$ (see Ebin and Marsden, 1970, Theorem 9.6).

[1] Readers unfamiliar with integration on manifolds may safely replace M in the following with $\mathbb{S}^1$. However, this does not simplify any of the argument.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

We have seen above that for vector fields solving Euler's equations, the associated flow yields a curve into the group of volume-preserving diffeomorphisms. Since g^{L^2} induces a weak Riemannian metric on $\mathrm{Diff}_\mu(M)$ which admits a metric spray, we can compute geodesics as curves which extremise the energy. This allows us to derive a differential equation for the flow corresponding to vector field solutions of (7.6). As in the Burgers case, we need that $g^{L^2}(h, \frac{\nabla}{\partial t}\frac{\partial}{\partial t}\eta)$ vanishes for every $h \in T\,\mathrm{Diff}_\mu(M)$. We know that the tangent space of $\mathrm{Diff}_\mu(M)$ is (up to a shift) given by divergence-free vector fields. Now due to the Helmholtz decomposition, Proposition E.17, elements which are L^2-orthogonal for every h are gradients of functions. Thus if $\frac{\nabla}{\partial t}\frac{\partial}{\partial t}\eta$ is a gradient then the inner product vanishes and the curve η extremises the energy. Conversely, if we assume that u solves (7.6) and denote by η its flow, we compute the derivative as follows:

$$\frac{\nabla}{\partial t}\frac{\partial}{\partial t}\eta(t,m) = ((\partial_t u) + \nabla_u u))(t,\eta(t)) = -\operatorname{grad} p(t,\eta(t,m)). \tag{7.9}$$

In other words, the Helmholtz decomposition shows that a flow η extremises the energy if and only if its associated vector field solves Euler's equation (7.6). Hence the Euler equation can equivalently be formulated as the following set of differential equations on the volume-preserving diffeomorphisms.

7.10 (Lagrangian formulation of the Euler problem) Find $\eta(t,\cdot) \in \mathrm{Diff}(M)$ for all t on some interval containing 0 such that

$$\begin{cases} \frac{\nabla}{\partial t}\frac{\partial}{\partial t}\eta(t,x) &= -\operatorname{grad} p(t,\eta(t,x)), \\ \eta(t,\cdot)_*\mu &= \mu \qquad \text{for all } t, \\ \eta(0,\cdot) &= \mathrm{id}_M\,. \end{cases} \tag{7.10}$$

These equations, (7.10), are called the *Euler equations in Lagrangian form*.

We achieved our goal to rewrite Euler's equation as a differential equation on an infinite-dimensional manifold. However, we have not yet exploited that the metric and the equation are right invariant (with respect to the group multiplication). In the next section we will investigate these properties and connect the Lagrangian formulation to the geometry of the Lie group at hand. This will lead (among other things) to another derivation of the geodesic equation as the Euler equations. While this might on first sight look like a superfluous exercise (after all we already know that Euler's equations can be rewritten as the geodesic equation) we wish to point out that this property is crucial for the investigation of PDEs in the Euler–Arnold framework we present here.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

7.3 Euler–Poincaré Equations on a Lie Group

Comparing the Lagrangian version of Euler's equations (7.10) and the L^2-metric, it is immediate that all terms arise by right-shifting objects. Moreover, as the tangent of the right shift with a diffeomorphism is just precomposition with the diffeomorphism, we obtain for a curve c with values in the volume-preserving diffeomorphisms, the formula

$$\mathrm{En}(c) = \int_0^1 \int_M g_m(\dot{c}(t)\circ c(t)^{-1}(m), \dot{c}(t)\circ c(t)^{-1}(m))\mathrm{d}\mu(m)\mathrm{d}t. \qquad (7.11)$$

Hence we can compute the energy using the L^2-inner product on the Lie algebra. We shall see that the geometry of the Lie group is tightly connected to the geodesics by virtue of the Riemannian metric being right invariant. To state Euler's equations as a geodesic equation, we need to understand derivatives of right-shifted variations.

7.11 Lemma *Let G be a Lie group and $p\colon]-\delta,\delta[\times[0,1] \to G$ a smooth variation. We identify $TG \cong G\times\mathbf{L}(G)$ by right multiplication (see Lemma 3.12) and define*

$$X_p\colon]-\delta,\delta[\times[0,1] \to \mathbf{L}(G), \quad (s,t) \mapsto \mathrm{pr}_2\left(T\rho_{p(s,t)^{-1}}\frac{\partial}{\partial t}p(s,t)\right),$$
$$Y_p\colon]-\delta,\delta[\times[0,1] \to \mathbf{L}(G), \quad (s,t) \mapsto \mathrm{pr}_2\left(T\rho_{p(s,t)^{-1}}\frac{\partial}{\partial s}p(s,t)\right).$$

Then the mixed derivatives of the right-shifted variations *are related as follows:*

$$\frac{\partial}{\partial s}X_p - \frac{\partial}{\partial t}Y_p = -[X_p, Y_p]. \qquad (7.12)$$

Proof It suffices to establish (7.12) pointwise for every pair (s_0,t_0). Define the smooth map $\tilde{p}(s,t) := p(s,t)\cdot p(s_0,t_0)^{-1}$. Then $\tilde{p}(s_0,t_0) = \mathbf{1}_G$ and a quick calculation shows that we have $\mathrm{pr}_2\left(T\rho_{\tilde{p}(s,t)^{-1}}\left(\frac{\partial}{\partial t}\tilde{p}(s,t)\right)\right) = X_p(s,t)$. A similar identity holds for Y_p and we may thus assume without loss of generality for the proof that $p(s_0,t_0) = \mathbf{1}_G$.

We work locally and pick a chart $\varphi\colon G \supseteq U \to V \subseteq \mathbf{L}(G)$ such that $\varphi(\mathbf{1}_G) = 0$ and $T_{\mathbf{1}_G}\varphi = \mathrm{id}_{\mathbf{L}(G)}$. As in 3.21 we define a local multiplication $v * w := \varphi(\varphi^{-1}(v)\varphi^{-1}(w))$ for all elements $v,w \in \mathbf{L}(G)$ near enough to zero. For an element v, we define (if it is close enough to 0) the inverse $v^{-1} := \varphi(\varphi^{-1}(v)^{-1})$. Choose an open 0-neighbourhood Ω such that $*$ is defined on $\Omega\times\Omega$ and Ω is symmetric (i.e. $v \in \Omega$ implies $v^{-1} \in \Omega$).

By construction, we have for all (s,t) in a neighbourhood of (s_0,t_0) that $q := \varphi\circ p$ makes sense and takes values in Ω. Note that also $q^{-1} = \varphi(p^{-1})$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

takes values in Ω for all such (s,t). Employing the rule on partial differentials, Proposition 1.20 shows that $X_p(s,t) = d_1 * \left(q(s,t), q^{-1}(s,t); \frac{\partial}{\partial t} q(s,t)\right)$. Specialising to (s_0,t_0), we see that $X_p(s_0,t_0) = \frac{\partial}{\partial t} q(s_0,t_0)$. Similar identities hold for Y_p by exchanging t and s. Compute the second derivative using the rule on partial differentials twice (where again the situation is symmetric in s and t):

$$\begin{aligned}\frac{\partial}{\partial s} X_p &= \frac{\partial}{\partial s} d_1 * \left(q, q^{-1}; \frac{\partial}{\partial t} q\right) \\ &= d_1^2 * \left(q, q^{-1}; \frac{\partial}{\partial t} q, \frac{\partial}{\partial s} q\right) + d_2(d_1 *) \left(q, q^{-1}; \frac{\partial}{\partial t} q, \frac{\partial}{\partial s} q^{-1}\right) \\ &\quad + d_1 * \left(q, q^{-1}; \frac{\partial^2}{\partial t \partial s} q\right).\end{aligned}$$

Due to Schwarz' rule, we see that the first and third terms in the above formula are completely symmetric in t and s. Hence these terms will not contribute to the difference (7.12). Let us now compute the differential of the inverse:

$$\begin{aligned}\frac{\partial}{\partial s} q^{-1}(s,t) &= \mathrm{pr}_2\, T\varphi \left(\frac{\partial}{\partial s} p(s,t)^{-1}\right) \\ &\overset{(3.3)}{=} -\mathrm{pr}_2\, T\varphi T\lambda_{p(s,t)^{-1}} T\rho_{p(s,t)^{-1}} \left(\frac{\partial}{\partial s} p(s,t)\right). \qquad (7.13)\end{aligned}$$

In particular, (7.13) reduces for (s_0,t_0) to $Y_p(s_0,t_0)$. Likewise for $\frac{\partial}{\partial t} q^{-1}(s_0,t_0)$ we obtain $X_p(s_0,t_0)$. Note that by construction $q(s_0,t_0) = 0 = q^{-1}(s_0,t_0)$. Hence we can now deduce that the difference $\left(\frac{\partial}{\partial s} X_p - \frac{\partial}{\partial t} Y_p\right)(s_0,t_0)$ is given as

$$\begin{aligned}&\left(d_2(d_1 *) \left(q, q^{-1}; \frac{\partial}{\partial t} q, \frac{\partial}{\partial s} q^{-1}\right) - d_2(d_1 *) \left(q, q^{-1}; \frac{\partial}{\partial s} q, \frac{\partial}{\partial t} q^{-1}\right)\right)(s_0,t_0) \\ &\overset{(7.13)}{=} -\Big(d_2(d_1 *) \left(0,0; X_p(s_0,t_0), Y_p(s_0,t_0)\right) \\ &\qquad + d_2(d_1 *) \left(0,0; Y_p(s_0,t_0), X_p(s_0,t_0)\right)\Big) \\ &= \frac{\partial^2}{\partial s \partial t}\bigg|_{s,t=0} \left(tY_p(s_0,t_0) * sX_p(s_0,t_0) - sX_p(s_0,t_0) * tY_p(s_0,t_0)\right) \\ &\overset{(3.21)}{=} -[X_p(s_0,t_0), Y_p(s_0,t_0)].\end{aligned}$$

□

We are now in a position to establish Arnold's classical result on the Euler equation via geometry on the Lie group of volume-preserving diffeomorphisms.

7.12 Theorem (Arnold) *Let (M,g) be a compact Riemannian manifold and consider a curve $\varphi\colon [0,1] \to \mathrm{Diff}_\mu(M)$. Then φ is a geodesic of the L^2-metric (i.e. the restriction of (7.8) to $\mathrm{Diff}_\mu(M)$) if and only if*

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$u := \dot{\varphi} \circ \varphi^{-1} \in \mathcal{V}_\mu(M)$$

solves the Euler equations (7.6)*, that is, for some function* $p\colon [0,1] \times M \to \mathbb{R}$*, we have*

$$\partial_t u + \nabla_u u = -\operatorname{grad} p.$$

Proof Let $\varphi(s,t)$ be a smooth variation of φ and $u(s,t) := \frac{\partial}{\partial t}\varphi(s,t)\circ\varphi(s,t)^{-1}$, that is, $u(t) = u(0,t)$. Set $h(t) = \frac{\partial}{\partial s}\varphi(s,t) \circ \varphi(s,t)^{-1}\Big|_{s=0}$ and note that by picking different smooth variations, $h(t)$ can be chosen to be an arbitrary curve in $\mathcal{V}_\mu(M)$ with vanishing endpoints. We now take the derivative of the energy and compute with (7.11):

$$\begin{aligned}\frac{d}{ds}\Big|_{s=0} \operatorname{En}(\varphi(s,\cdot)) &= \int_0^1 \int_M g_m\left(u(t)(m), \frac{\partial}{\partial s}\Big|_{s=0} u(s,t)(m)\right) \mathrm{d}\mu(m)\mathrm{d}t \\ &\overset{(7.12)}{=} \int_0^1 g^{L^2}\left(u(t), \frac{d}{dt}h(t) - [u(t), h(t)]\right) \mathrm{d}t. \end{aligned} \tag{7.14}$$

Recall that the Lie bracket in (7.14) is the Lie bracket of $\mathcal{V}_\mu(M)$. As this is a subalgebra of the Lie algebra of $\operatorname{Diff}(M)$, Example 3.25 implies that it is the negative of the commutator bracket of vector fields, whence $\nabla_u h - \nabla_h u = -[u,h]$. Now replace the Lie bracket and apply Exercise E.3.5 to $g(u, -\nabla_h u)$ to see that (7.14) yields

$$\begin{aligned}&\int_0^1 g^{L^2}\left(u(t), \frac{d}{dt}h(t) + \nabla_{u(t)}h(t) - \nabla_{h(t)}u(t)\right)\mathrm{d}t \\ &= \int_0^1 \Bigg(g^{L^2}\left(u(t), \frac{d}{dt}h(t) + \nabla_{u(t)}h(t)\right) \underbrace{-\frac{1}{2} g^{L^2}(\operatorname{grad} g(u(t),u(t)), h(t))}_{=0 \text{ by Prop. E.17 as } h(t) \in \mathcal{V}_\mu(M)} \Bigg) \mathrm{d}t \\ &= \int_0^1 g^{L^2}\left(u(t), \frac{d}{dt}h(t) + \nabla_{u(t)}h(t)\right)\mathrm{d}t. \end{aligned}$$

We continue with integration by parts with respect to t and the identity $g(u, \nabla_u h) = g(u, \operatorname{grad} g(u,h)) - g(\nabla_u u, h)$ (see Exercise E.3.5). Together with the Helmholtz decomposition, Proposition E.17, the above equation is equal to

$$-\int_0^1 g^{L^2}\left(\frac{d}{dt}u(t) + \nabla_{u(t)}u(t), h(t)\right)\mathrm{d}t.$$

Hence if φ extremises the energy, we see that $-(\partial_t u + \nabla_u u)$ must be L^2-orthogonal to every curve (with vanishing endpoints) in $\mathcal{V}_\mu(M)$. By the Helmholtz decomposition, this happens if and only if there is some function p (determined up to a constant) such that $\partial_t u + \nabla_u u = -\operatorname{grad} p$. □

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

We have now seen that the geometry of the group of volume-preserving diffeomorphisms can be exploited to identify the Euler equation as a geodesic equation. This connection is typical for PDEs and their associated geodesic equations which are amenable to Arnold's approach. Indeed there is one last reformulation of the Euler equation on the Lie group $\mathrm{Diff}_\mu(M)$ which needs to be mentioned here as it exhibits the connection between the invariant Riemannian metric and Lie group more explicitly.

7.13 Remark (The Euler equation as an Euler–Poincaré equation on $\mathrm{Diff}_\mu(M)$) Our aim is to identify the geodesic equation as a so-called *Euler–Poincaré equation* on $\mathrm{Diff}_\mu(M)$. For this, let us start more generally: Let G be a regular Lie group with Lie algebra $(\mathbf{L}(G), [\cdot,\cdot])$ and $\langle\cdot,\cdot\rangle$ a continuous inner product on $\mathbf{L}(G)$. Assume that we wish to compute geodesics for the right-invariant metric induced by the choice of inner product. Arguing as for the Euler equation, we see that a curve $\varpi\colon [0,1] \to G$ extremises the energy if and only if the expression (7.14) vanishes, that is, in the notation of Exercise 3.2.11 we must have

$$0 = \int_0^1 \left\langle u(t), \frac{d}{dt}h(t) - [u(t), h(t)] \right\rangle \mathrm{d}t = \int_0^1 \left\langle u(t), \frac{d}{dt}h(t) - \mathrm{ad}_{u(t)}(h(t)) \right\rangle \mathrm{d}t,$$

where $u(t) = \delta^r\varphi(t)$ (the right logarithmic derivative of φ). Now assume that for all $x \in \mathbf{L}(G)$, there exists an adjoint $\mathrm{ad}_x^\top$ for the linear operator ad_x with respect to the inner product $\langle\cdot,\cdot\rangle$, that is, $\langle \mathrm{ad}_x^\top(y), z\rangle = \langle y, \mathrm{ad}_x(z)\rangle$. Applying again integration by parts we see that φ is a geodesic if and only if its right logarithmic derivative satisfies the *Euler–Poincaré equation*[2]

$$\frac{d}{dt}\delta^r\varphi = -\mathrm{ad}^\top_{\delta^r\varphi}(\delta^r\varphi).$$

Thus we have derived yet another expression which is equivalent to the geodesic equation and by the previous results also to the Euler equation of an incompressible fluid if $G = \mathrm{Diff}_\mu(M)$ and the inner product is the L^2-inner product. Observe that the Euler–Poincaré equation is a differential equation on the Lie algebra. The Euler–Poincare equation reduces the geodesic equation to the Lie algebra and shows that the geometry of the Lie group and the Riemannian geometry of a right-invariant metric are closely intertwined. We will not discuss this fruitful perspective on the Euler equation. However, there are accounts of the general mechanism with many examples available in the literature. The interested reader is referred, for example, to Vizman (2008), Modin (2019) and Khesin and Wendt (2009, II.3).

[2] The name goes back to honor Henri Poincaré who formulated in Poincaré (1901) differential equations for mechanical systems on (finite-dimensional) Lie groups in the presented form.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Exercises

7.3.1 Let (M,g) be a compact Riemannian manifold. Show that the L^2-metric (7.8) on $\mathrm{Diff}(M)$ restricts to a right-invariant Riemannian metric on $\mathrm{Diff}_\mu(M)$.

7.3.2 We again let (M,g) be a compact Riemannian manifold and consider the L^2-metric g^{L^2} (7.8) on $\mathrm{Diff}(M)$. Let S be the metric spray of g and K the associated connector. The aim of this exercise is to prove that g^{L^2} admits a metric spray, connector and covariant derivative by exploiting the right invariance of the metric.
Remark: The proof for this statement for the L^2-metric on $C^\infty(\mathbb{S}^1,M)$ from §5.1 can be adapted to the present situation. However, we will follow in this exercise the classical argument of Ebin and Marsden (1970) which highlights the use of invariance properties.

(a) Show that the pushforward S_* is a spray on $\mathrm{Diff}(M)$ and the pushforward K_* is a connector on $\mathrm{Diff}(M)$.

(b) Define the covariant derivative $\nabla^{L^2}_X Y = K_* \circ TY \circ X$ associated to the connector K (i.e. X,Y are vector fields on $\mathrm{Diff}(M)$). Work out $\nabla^{L^2}_X Y$ for right-invariant vector fields on $\mathrm{Diff}(M)$ (i.e. vector fields $X(\varphi) = X(\mathrm{id}) \circ \varphi$, where $X(\mathrm{id}) \in \mathcal{V}(M) = \mathbf{L}(\mathrm{Diff}(M))$. Then verify that ∇^{L^2} satisfies the properties of the metric derivative associated to g^{L^2} for all right-invariant vector fields.

(c) Establish that ∇^{L^2} is the metric derivative of g^{L^2} and deduce that K_* is the connector and S_* the metric spray associated to g^{L^2}.
Hint: Exploit the idea that for every vector field $X \in \mathcal{V}(\mathrm{Diff}(M))$ and $\varphi \in \mathrm{Diff}(M)$ there exists a right-invariant vector field X^R such that $X^R(\varphi) = X(\varphi)$.

7.3.3 Supply the necessary details for the proof of Lemma 7.11. Show in particular that $\mathrm{pr}_2\left(T\rho_{\tilde{p}(s,t)^{-1}}\left(\frac{\partial}{\partial s}\tilde{p}(s,t)\right)\right) = X_p(s,t)$, $X_p = d_1 *(q,q^{-1};\frac{\partial}{\partial s}q)$ and $X_p(s_0,t_0) = \frac{\partial}{\partial s}q(s_0,t_0)$ and (7.13) reduces to $X_p(s_0,t_0)$.

7.4 An Outlook on Euler–Arnold Theory

In the last section we saw how the Euler equations of an incompressible fluid can be identified as a geodesic equation of a right-invariant metric on an infinite-dimensional Lie group. This equation can in turn be rephrased as the Euler–Poincaré equation on the Lie algebra, highlighting the close connection of Lie group and Riemannian geometry. In the present section we will discuss several

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

applications of the theory developed so far. The section concludes with a discussion of the problems one faces related to carrying out the program sketched.

7.14 (Applications of Euler–Arnold theory) As we have seen, certain PDEs, such as the Euler equation of an incompressible fluid, can be rewritten as geodesic equations on an infinite-dimensional manifold. Moreover, the mechanism is reversible, that is, solutions to the original PDE correspond to geodesics. Vice versa, solutions to the geodesic equation of certain (weak) Riemannian metrics yield solutions to partial differential equations. This has the following immediate applications.

(a) *Existence, uniqueness and parameter dependence of solutions*. Geodesics are solutions to ordinary differential equations. To establish properties of their solutions (such as local existence of unique solutions) one can hope to apply the usual toolbox for ordinary differential equations. Transporting solutions back to the finite-dimensional world, this will yield local existence and uniqueness for solutions of the PDE. Historically, this was how the existence and uniqueness problem for Euler's equations of an incompressible fluid was first solved in the general case in Ebin and Marsden (1970).[3] The caveat here is that ODE tools break down beyond Banach manifolds and one requires a technical analysis to make the program work (see 7.18).

Before we continue let us recall the concept of sectional curvature.

7.15 Let (M,g) be a (weak) Riemannian manifold with covariant derivative ∇ and curvature R. Then for two linearly independent vectors $u,v \in T_m M$ spanning a 2-dimensional subspace σ the *sectional curvature* is defined as

$$K(\sigma) := \frac{g_x(R(v,w)w,v)}{g_x(v,v)g_x(w,w) - (g_x(v,w))^2}$$

(where we actually evaluate the curvature R in (local) vector fields V,W with $V(x) = v$ and $W(x) = w$).

As a concrete example, endow the diffeomorphism group $\mathrm{Diff}(M)$ with the weak L^2-metric. Every vector $V \in T_\varphi \mathrm{Diff}(M)$ can be expressed as the value of a right-invariant vector field X_V on $\mathrm{Diff}(M)$. Hence it suffices to compute the sectional curvature using right-invariant vector fields. The formula for the covariant derivative of the L^2-metric then shows that for a subspace σ

[3] The emphasis here is on 'general'. Some results were known for special cases previous to the treatment in loc. cit.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

generated by the orthonormal elements $\{X_V(\varphi), Y_V(\varphi)\}$, the sectional curvature is given as

$$K^{L^2}(\sigma) = g^{L^2}(R^{L^2}(X_V, Y_V)Y_V, X_V) = \int_M K(X_V(\mathrm{id})(m), Y_V(\mathrm{id})(m))\mathrm{d}\varphi^*(\mu)(m),$$

where K is the sectional curvature on M (computed with respect to the space spanned by the vectors $X_V(\mathrm{id})(m)$ and $Y_V(\mathrm{id})(m)$; Smolentsev, 2007, 6.4).

Continuing with our review of Euler–Arnold theory from 7.14:

(b) *Geometric tools for PDE analysis*. It is well known from finite-dimensional Riemannian geometry that curvature controls the behaviour of geodesics (see e.g. do Carmo, 1992, Chapter 5, and also note the connection to the Hopf–Rinow theorem in Remark 4.44). The point is that for positive sectional curvature, geodesics starting at the same point with slight variation of the initial velocity tend to converge towards each other, while for negative sectional curvature they diverge (this can be made explicit as in the finite-dimensional case, but we will not discuss the details here). In the context of partial differential equations these properties can be interpreted as stability of solutions under perturbations of initial conditions.

Indeed Arnold (1966) showed that the sectional curvature of the L^2-metric on $\mathrm{Diff}_\mu(M)$ is negative in almost all directions. So nearby fluid regions will typically diverge exponentially fast from each other. This analysis applies, in particular, to partial differential equations employed in weather forecasts. So infinite-dimensional Riemannian geometry shows that reliable long-term weather forecasts are practically impossible.

7.16 Remark There is also a beautiful connection of the Euler–Arnold equations to ideas from Hamiltonian mechanics on $\mathrm{Diff}(M)$. The differential geometric context for this is (weakly) symplectic structures on $\mathrm{Diff}(M)$ and we refer to Smolentsev (2007, §§6 and 7) for a discussion.

In the present chapter we have only seen the mechanism applied to the Burgers and Euler equations. There are many more PDEs which typically arise in hydrodynamics and can be treated in the same framework. Equations which are amenable to this treatment are nowadays called *Euler–Arnold equations*. We refer to Khesin and Wendt (2009) for an extensive list but mention explicitly the Camassa–Holm equation, the Hunter–Saxton equation and the Korteweg–deVries (KdV) equation as PDEs belonging to this class. Since the Hunter–Saxton equation admits a beautiful geometric interpretation, we will now briefly discuss a few more details related to this equation and its geometric treatment.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

7.17 Example (Hunter–Saxton equation) We consider the *periodic Hunter–Saxton equation* on the circle $\mathbb{S}^1$. The task is to find a time-dependent vector field $u\colon [0,T[\times\mathbb{S}^1 \to \mathbb{R}$ which satisfies the following equation:

$$u_{txx} + 2u_x u_{xx} + uu_{xxx} = 0, \tag{7.15}$$

where subscripts again denote partial derivatives with respect to time (t) or $x \in \mathbb{S}^1$. We will see in Exercise 7.4.1 that the Hunter–Saxton equation is the geodesic equation of the right-invariant $\dot{H}^1$-semimetric. To describe this semimetric, we recall from §5.1 the notation $U' := T_\theta U(1)$ and define the inner product

$$g^{\dot{H}^1}_{\mathrm{id}}(U,V) = \frac{1}{4}\int_{\mathbb{S}^1} U'(\theta)V'(\theta)\mathrm{d}\theta \qquad \text{on } C^\infty(\mathbb{S}^1,\mathbb{R}). \tag{7.16}$$

Then the $\dot{H}^1$-semimetric is the right-invariant semimetric induced by (7.16). Note that it is a semimetric as constant vector fields are annihilated. In Example 3.42 we saw that the constant vector fields generate the group of rotations $\mathrm{Rot}(\mathbb{S}^1)$. Hence there are two possibilities to obtain a (weak) Riemannian metric: One can work with the quotient manifold $\mathrm{Diff}(\mathbb{S}^1)/\mathrm{Rot}(\mathbb{S}^1)$ (see Lenells, 2008 for a detailed discussion) or one has to fix a subgroup containing only the trivial rotation. We consider the induced weak Riemannian metric on the Lie subgroup

$$D_0 := \{\phi \in \mathrm{Diff}(\mathbb{S}^1) \mid \phi(\theta(0)) = \theta(0)\}, \tag{7.17}$$

where $\theta\colon [0,2\pi] \to \mathbb{R}^2$, $t \mapsto (\cos(t),\sin(t))$ is the canonical parametrisation of the circle. To ease the computations, we follow Lenells (2008) and will in the following always identify diffeomorphisms and vector fields of $\mathbb{S}^1$ as periodic mappings $[0,2\pi] \to \mathbb{R}$. This allows one to prove that the Hunter–Saxton equation exhibits a fascinating geometric feature discovered in Lenells (2007): The group D_0 can be identified as a convex subset of a sphere and this embedding is a Riemannian isometry relating the $\dot{H}^1$-metric to the L^2-metric. Indeed this embedding is surprisingly simple, as it is given by

$$\Psi\colon D_0 \to C^\infty(\mathbb{S}^1,\mathbb{R}), \quad \Psi(\varphi) = \sqrt{\varphi'}.$$

Its image becomes the convex set

$$\{f(\theta) > 0, \text{ for all } \theta \in \mathbb{S}^1, \int_{\mathbb{S}^1} |f(\theta)|^2\mathrm{d}\theta = 1\} \subseteq C^\infty(\mathbb{S}^1,\mathbb{R}).$$

We omit the details here and refer instead to the exposition in Lenells (2007).

The geometric content of this observation is that we can transform the Hunter–Saxton equation to a geodesic equation on the L^2-sphere in $C^\infty(\mathbb{S}^1,\mathbb{R})$. Now solutions to the geodesic equation on the L^2-sphere (with respect to the natural L^2-metric) can be explicitly computed: We have seen in Example 4.43

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

that geodesics of the L^2-sphere (seen as the unit sphere of the Hilbert space $L^2(\mathbb{S}^1,\mathbb{R})$) are given by great circles.

While the Hunter–Saxton equation can be interpreted as the geodesic equation on an infinite-dimensional sphere in a Hilbert space, our approach so far has been to consider this equation as an equation on the space $C^\infty(\mathbb{S}^1,\mathbb{R})$ which is not a Hilbert space. This is quite unnatural for two reasons: From the perspective of the PDEs this approach will only allow solutions which are smooth in space, whereas it is often of interest to have much less regular solutions, for example, solutions which are only finitely often differentiable in space. The geometric perspective allows us to connect the PDE to an ODE which we then have to solve. However, on $\mathrm{Diff}(M)$ this presents a problem, as we will discuss now.

7.18 (Returning to the Banach and Hilbert setting) In 7.14 we listed as an advantage of the Euler–Arnold approach to PDEs that (local) existence and uniqueness of solutions to these PDEs can be obtained by methods for ordinary differential equations (ODEs) on infinite-dimensional manifolds. Unfortunately the manifolds we have been working in this chapter are submanifolds of the manifolds of mappings $C^\infty(K,M)$ (where K is a compact manifold). These manifolds are never (except in trivial cases) Banach manifolds, so there are no black-box techniques for ODEs as Appendix A.6 shows. Thus the elegant theory developed so far misses an essential analytic ingredient to solve the ODEs occurring.

The solution to this problem is, in principle, simple (if one glosses over the technical details): Replace C^∞ functions by finitely often differentiable ones. Note that this dovetails nicely with the problem statement from the PDE side we mentioned earlier. By going to finitely often differentiable functions, we allow solutions to the PDE which are much less regular in space. From the perspective of infinite-dimensional manifolds, one can prove that

$$C^\infty(K,M) = \bigcap_{k\in\mathbb{N}_0} C^k(K,M) = \bigcap_{s\in\mathbb{N}_0} H^s(K,M),$$

where the manifolds in the middle are Banach manifolds and the spaces on the right even Hilbert manifolds. Here $H^s(K,M)$ denotes all mappings of Sobolev H^s-type (meaning that their weak derivatives are in L^2). We refrain from defining Sobolev spaces on manifolds as there are several subtle points involved in their construction. Instead we remark that they admit a manifold structure similar to the manifolds of mappings we constructed in Chapter 2. For more information we refer the reader to the detailed exposition in Inci et al. (2013).

Conveniently, Sobolev type groups of diffeomorphisms $\mathrm{Diff}^{H^s}(K)$ also exist and one can even prove that $\mathrm{Diff}(K) = \lim_{s\to\infty} \mathrm{Diff}^{H^s}(K)$ as a projective

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

limit in the category of manifolds. This structure is called an ILH–Lie group (see Omori, 1974), and coincides with the Lie group structure of $\mathrm{Diff}(K)$ constructed in Example 3.5. The point of the construction is of course that one can work on the Hilbert manifold of Sobolev morphisms. Unfortunately, because of Omori's theorem (Omori, 1978), the groups $\mathrm{Diff}^{H^s}(K)$ cannot be Lie groups. They are, however, manifolds and topological groups such that the left multiplication is continuous but not differentiable – see (2.6) – to see that the derivative loses orders of differentiability. However, right multiplication is still smooth and one obtains a so-called *half-Lie group* (Marquis and Neeb, 2018). This leads to several analytic problems which need to be solved to establish smoothness of the associated metric sprays (see e.g. Ebin and Marsden, 1970). We refer the reader to the literature for more details as these problems are beyond the scope of this chapter.

By this point, the reader should be suitably equipped to understand the classical research literature on these topics. For example, Ebin and Marsden (1970) as well as Ebin (2015) present the theory for the Euler equation of an incompressible fluid. Also the monograph of Khesin and Wendt (2009) provides an excellent overview of the theory together with many pointers towards the literature. This chapter concludes with a short remark on some more recent developments in Euler–Arnold theory.

7.19 Remark Euler–Arnold theory is still an active area of research. Among the many recent results, I like to point out several which I find particularly interesting:

(a) Classical Euler–Arnold theory works with right-invariant Riemannian metrics on (subgroups of) diffeomorphism groups. In Bauer and Modin (2020) it was shown that the approach also works for Riemannian metrics on $\mathrm{Diff}(M)$ which are only invariant with respect to $\mathrm{Diff}_\mu(M)$. Thus a whole new family of PDEs, such as certain shallow water equations, can be treated by Euler–Arnold methods.

(b) Instead of a purely deterministic PDE one can apply the mechanism to a stochastic partial differential equation. Stochastic versions of Euler's equations for an incompressible fluid have recently been considered as models for data-driven hydrodynamics (modelling uncertainty in the data). For the Euler equation of an incompressible fluid, Maurelli et al. (2019) work out the necessary details to make the Euler–Arnold machinery work in the stochastic setting.

(c) There is a connection between Euler–Arnold theory, the differential geometry of diffeomorphism groups and optimal mass transport. This is based

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

on the observation that there is a submersion $\pi\colon \mathrm{Diff}(M) \to \mathrm{Dens}(M)$ from the diffeomorphism group to the manifold of densities on M. This links the Riemannian metrics on $\mathrm{Diff}(M)$ to the Wasserstein metric from optimal transport. An introduction to these topics can be found in Khesin and Wendt (2009, Appendix A.5).

Exercises

7.4.1 We consider the group $\mathrm{Diff}(\mathbb{S}^1)$ and the $\dot{H}^1$-semimetric (7.16). Show that:

(a) Identifying diffeomorphisms with periodic mappings we can identify D_0 with

$$\{u + \mathrm{id} \mid u\colon [0,2\pi] \to \mathbb{R}, \quad \frac{d}{dx}u > -1, u(0) = 0 = u(2\pi)\}.$$

(b) $E = \{u \in C^\infty([0,2\pi],\mathbb{R}) \mid u(0) = 0\}$ is a closed subspace of $C^\infty([0,2\pi],\mathbb{R})$ and D_0 is diffeomorphic to an open subset of $\mathrm{id} + E \subseteq E$.

(c) The continuous linear operator $A(u) := -u''$ induces an isomorphism from E to $F := \{f \in C^\infty([0,2\pi],\mathbb{R}) \mid \int_0^{2\pi} f(x)\mathrm{d}x = 0\}$. Moreover, show that (up to identification) $g^{\dot{H}^1}_{\mathrm{id}}(U,V) = \int_{\mathbb{S}^1} UA(V)\mathrm{d}\theta$.

(d) The group D_0 from (7.17) is a Lie subgroup of $\mathrm{Diff}(\mathbb{S}^1)$ and its Lie algebra can be identified as

$$\mathbf{L}(D_0) = \{f \in C^\infty(\mathbb{S}^1,\mathbb{R}) \mid \int_{\mathbb{S}^1} f(\theta)\mathrm{d}\theta = 0\}.$$

(e) (7.16) induces a right-invariant weak Riemannian metric on D_0.

(f) A curve φ extremises the energy of the $\dot{H}^1$-semimetric (7.16) if and only if $u = \frac{\partial}{\partial t}\varphi \circ \varphi^{-1}$ satisfies the Hunter–Saxton equation (7.15).

7.4.2 Prove the claims on the sectional curvature of the L^2-metric on $\mathrm{Diff}(M)$ from (7.15).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

8

The Geometry of Rough Paths

8.1 Introduction

In this chapter we will discuss the (infinite-dimensional) geometric framework for rough paths and their signature. Rough path theory originated in the 1990s with the work of T. Lyons; see, for example, Lyons (1998). It seeks to establish a theory of integrals and differential equations driven by rough signals. For example, one is interested in the controlled ordinary differential equations of the following type:

$$y_t' = f(y_t) + g(y_t)X_t'. \tag{8.1}$$

Here the subscripts track the time parameter t, X is an input path with values in $\mathbb{R}^d$ and y is the output with values in $\mathbb{R}^e$. We will for this exposition assume that all derivatives and integrals needed exist (e.g. if X is a smooth path). Finally, f, g are non-linear functions with values in $\mathbb{R}^e$ and $L(\mathbb{R}^d, \mathbb{R}^e)$, i.e. $e \times d$-matrices, respectively. Focussing on the control term in (8.1), let us consider a simple approximation to the solution in the case that $f = 0$, that is, $y_t' = g(y_t)X_t'$. For example, the first-order Euler method gives us the following approximation for the components of y:

$$y_t^i - y_s^i \approx g^i(y_s) \int_s^t dX^i = \lim_{|P|\to 0} g^i(y_s) \sum_{[t_j, t_{j+1}]\in P} \left(X^i_{t_{j+1}} - X^i_{t_j}\right), \quad i = 1, \dots, e.$$

Here superscripts denote components of maps, the integral is defined via a Riemann–Stieltjes sum and we think of $g(y_s)$ as a matrix (selecting columns and rows appropriately). The information needed for the approximation is the integral of X. In general, we would like better approximations (or even a solution), so it is natural to increase the order of the approximations. To obtain the desired formula one applies a Taylor expansion to obtain the following second-order Euler approximation for $i = 1, \dots, e$:

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$y_t^i - y_s^i \approx g^i(y_s)\int_s^t dX^i + \sum_{k=1}^{e}\sum_{j=1}^{d} g_k^i(y_s)\frac{\partial}{\partial_k}g^j(y_s)\int_s^t\int_s^r \mathrm{d}X^i\mathrm{d}X^j. \quad (8.2)$$

Thus higher-order approximation requires knowledge about (mixed) iterated integrals of the path X against itself. So far we have tacitly assumed that the necessary paths and derivatives exist (e.g. that all the objects in question are smooth).

Weakening this requirement, the objects we are interested in are iterated integrals of Hölder continuous paths with values in $\mathbb{R}^d$.[1] For convenience we will assume throughout this chapter that we are dealing with paths on the interval $[0,1]$. This is no restriction since we can always reparametrise Hölder paths on any $[0,1]$ to obtain corresponding paths on $[0,1]$ (though this changes the Hölder norm). Assume we have two continuous mappings $X\colon [0,1]\to\mathbb{R}^d$ and $Y\colon [0,1]\to L(\mathbb{R}^d,\mathbb{R}^e)$, where the continuous linear maps $L(\mathbb{R}^d,\mathbb{R}^e)$ have been endowed with the operator norm. We would like to define an integral now of Y against the path X and these integrals should yield a continuous map $(X,Y)\mapsto\int Y\mathrm{d}X$. Setting $X_{s,t}\coloneqq X_t - X_s$, we could try to define the integral as a limit of Riemann–Stieltjes sums,

$$\int_0^1 Y(t)\mathrm{d}X(t) \coloneqq \lim_{|\mathcal{P}|\to 0}\sum_{[s,t]\in\mathcal{P}} Y(s)X_{s,t},$$

where $\mathcal{P}$ is a partition of $[0,1]$ and the limit takes the mesh size to 0. The resulting integral is called the *Young integral* and Young (1936) showed that the Riemann–Stieltjes sum converges if X is an α-Hölder path[2] and Y is a β-Hölder path such that $\alpha+\beta>1$. This result is sharp as one can construct examples of paths with $\alpha+\beta=1$ such that the sum becomes ill defined. So in general, there is no hope for an integration theory which allows us to integrate arbitrarily rough paths against each other. However, the key insight of rough path theory is that the regularity assumption $\alpha+\beta>1$ from Young's theorem can be circumvented if additional structure is added to the paths. Thus a rough integral can be built if we enhance Hölder continuous paths with additional information to so-called rough paths. The point is that this information can be chosen for irregular paths such as Brownian motion (which is known to be α-Hölder for $\alpha\in]0,1/2[$). Here rough path theory excels at clever estimates for the integrals appearing. In the present chapter we will focus on the geometric side of the picture and leave the hard analytic estimates to the rough path literature.

[1] The basic theory extends to Banach space valued paths but this requires more technical efforts such as, for example, the introduction of tensor norms.

[2] The notion of α-Hölder paths is recalled in §8.3.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Our first aim is to develop a convenient geometric framework to record iterated integrals.

Notation for increments and iterated integrals We frequently encounter increments of continuous paths $X\colon [0,1] \to \mathbb{R}^d$:

- $X_t := X(t)$, and
- $X_{s,t} := X_t - X_s$ for the increment and $t, s \in [0,1]$.

With enough differentiability of X, we can define iterated integrals of X against itself. For $0 \le s \le t \le 1$, we integrate componentwise in $\mathbb{R}^d \otimes \mathbb{R}^d \cong \mathbb{R}^{d^2}$ and set

$$\int_s^t \int_s^r \mathrm{d}X \otimes \mathrm{d}X := \int_s^t X_{s,r} \otimes \mathrm{d}X := \int_s^t X_{s,r} \otimes \mathrm{d}X_r := \int_s^t X_{s,r} \otimes X_r' \mathrm{d}r.$$

The notation suppresses indices with the understanding that the objects are matrices whose components are iterated integrals. For higher iterated integrals this will quickly become impractical whence we shall use tensor notation instead.

8.2 Iterated Integrals and the Tensor Algebra

In this section we consider the tensor algebra as a continuous inverse algebra. As seen in the introduction to this chapter, we are interested in iterated integrals of the components of a path $X\colon [0,1] \to \mathbb{R}^d$. For example, the second iterated integrals (8.2) yield a matrix object whose components can also be conveniently recorded using tensor notation:

$$\int_s^t \int_s^r \mathrm{d}X^i \otimes \mathrm{d}X^j := \int_s^t \left(\int_s^r \mathrm{d}X^i \right) \mathrm{d}X^j e_i \otimes e_j,$$

where e_i, e_j are standard basis vectors of $\mathbb{R}^d$. Iterating for higher orders, we can write the resulting integrals as elements in an iterated tensor product. Note that on the basic levels tensors are just a bookkeeping device to track the components integrated against each other. Indeed the canonical identification $\mathbb{R}^d \otimes \mathbb{R}^d \cong \mathbb{R}^{d \times d}$ maps $e_i \otimes e_j$ to the component in the ith row and jth column. Hence, detailed knowledge on tensor products is not needed and we refer to (Abraham et al., 1988, Chapter 5) for an introduction.

8.1 Definition ((Truncated) tensor algebra) For $d \in \mathbb{N}$ and $k \in \mathbb{N}$, we set

$$(\mathbb{R}^d)^{\otimes k} := \underbrace{\mathbb{R}^d \otimes \mathbb{R}^d \otimes \cdots \otimes \mathbb{R}^d}_{k \text{times}} \text{ and } (\mathbb{R}^d)^{\otimes 0} := \mathbb{R}.$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Write $e_i, 1 \le i \le d$ for the standard basis vectors of $\mathbb{R}^d$ and recall that the products $e_{i_1} \otimes \cdots \otimes e_{i_k}$ form a base for $(\mathbb{R}^d)^{\otimes_k}$ whence this space is isomorphic to $\mathbb{R}^{d^k}$. Then we define for $N \in \mathbb{N} \cup \{\infty\}$ the *(truncated) tensor algebra*[3]

$$\mathcal{T}^N(\mathbb{R}^d) := \prod_{k=0}^{N} (\mathbb{R}^d)^{\otimes k}.$$

Elements in $\mathcal{T}^N(\mathbb{R}^d)$ will be denoted as sequences $(x_k)_{k<N+1}$ (where $\infty + 1 = \infty$). An element concentrated in the kth factor $(\mathbb{R}^d)^{\otimes k}$ is called *homogeneous of degree* k. Then the algebra product is given by

$$(x_k)_{k<N+1} \otimes (y_k)_{k<N+1} := \left(\sum_{n+m=k} x_n \otimes y_m \right)_{k<N+1}. \tag{8.3}$$

Since tensor products over $\mathbb{R}$ with elements of $\mathbb{R}$ contract, that is, $\lambda \otimes v = \lambda v$, we see that the tensor algebra has a unit element $\mathbf{1} = (1,0,0,\ldots) \in \mathcal{T}^N(\mathbb{R}^d)$ (i.e. $\mathbf{1}$ is homogeneous of degree 0) and the map $\pi_0^N : \mathcal{T}^N(\mathbb{R}^d) \to (\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$ is an algebra morphism (and, in particular, the homogeneous elements of degree 0 form a subalgebra of the (truncated) tensor algebra). Note, however, that the algebra is almost always non-commutative (Exercise 8.2.2).

In the introduction we saw that iterated integrals can be conveniently identified with elements in the tensor algebra. From this point of view it would be enough to treat the tensor algebra as locally convex space which simply stores information. However, it turns out that iterated integrals satisfy several natural identities which can be expressed using the product in the tensor algebra. For example, if $X\colon [0,1] \to \mathbb{R}^d$ is a smooth path and $\mathbf{X}_{s,t} := \int_s^t X_{s,r} \otimes \mathrm{d}X \in \mathbb{R}^d \otimes \mathbb{R}^d$ its iterated integral, it is easy to see that these satisfy *Chen's relation*

$$\mathbf{X}_{s,t} - \mathbf{X}_{s,u} - \mathbf{X}_{u,t} = X_{s,u} \otimes X_{u,t}, \quad \text{for all } u \in [s,t]. \tag{8.4}$$

To get a feeling for these identities and for the truncated tensor algebra, we recommend Exercise 8.2.2. Summing up, we should be interested in the algebra structure of the tensor algebra as well. Our next result discusses the topological structure of these algebras.

8.2 Lemma *Fix $N \in \mathbb{N} \cup \{\infty\}$ and $d \in \mathbb{N}$.*

(a) *An element a in $\mathcal{T}^N(\mathbb{R}^d)$ is invertible in the tensor algebra if and only if the associated element of degree 0, $a_0 := \pi_0^N(a)$ is invertible.*

[3] The tensor algebra discussed in this section differs from what is usually called the tensor algebra. In the literature, the tensor algebra usually denotes the direct sum of iterated tensor products (i.e. finite sequences of iterated tensors). From this perspective, the tensor algebra $\mathcal{T}^\infty(\mathbb{R}^d)$ should rather be called the completed tensor algebra.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(b) *The algebra $\mathcal{T}^N(\mathbb{R}^d)$ is a continuous inverse algebra (CIA). Moreover, if $N < \infty$, then $\mathcal{T}^N(\mathbb{R}^d)$ is a Banach algebra and if $N = \infty$, $\mathcal{T}^N(\mathbb{R}^d)$ is a Fréchet algebra.*

Proof As a first step, let us topologise $\mathcal{T}^N(\mathbb{R}^d)$.

If $N < \infty$, we obtain a finite-dimensional algebra and thus it is a Banach algebra and a CIA (alternatively see Exercise 8.2.1).

If $N = \infty$, we observe first that $\mathcal{T}^\infty(\mathbb{R}^d)$ is a countable product of finite dimensional locally convex spaces, whence it is a Fréchet space. Note that this topology and the Banach topology for $N < \infty$ turn the projection π_0^N into a continuous map. Exploiting the product topology, we see that the product $\mathcal{T}^\infty(\mathbb{R}^d) \times \mathcal{T}^\infty(\mathbb{R}^d) \to \mathcal{T}^\infty(\mathbb{R}^d)$ is continuous if and only if the component maps $P_k : \mathcal{T}^\infty(\mathbb{R}^d) \times \mathcal{T}^\infty(\mathbb{R}^d) \to (\mathbb{R}^d)^{\otimes k}, k \in \mathbb{N}$ are continuous. However, as the algebra product (8.3) respects the homogeneous degree of elements, it is clear that P_k factors through the continuous inclusion $\mathcal{T}^k(\mathbb{R}^d)$ and the algebra product of $\mathcal{T}^k(\mathbb{R}^d)$. Both the inclusion and the algebra product are continuous, and so P_k and consequently the product on $\mathcal{T}^\infty(\mathbb{R}^d)$ are continuous. We deduce that $\mathcal{T}^\infty(\mathbb{R}^d)$ is a locally convex algebra. To see that it is also a CIA, let us prove first the claim on invertibility of elements.

Let $a \in \mathcal{T}^N(\mathbb{R}^d)$ and write $a = a_0 + b$, where $a_0 = \pi_0^N(a)$ and $b = a - a_0$. Since π_0^N is an algebra homomorphism, a_0 must be invertible if a is invertible. For the converse, we observe that a is invertible if and only if $a_0^{-1}a = \mathbf{1} + a_0^{-1}b$ is invertible. Plugging $a_0^{-1}a$ into the *Neumann inversion formula* (see Werner, 2000, Theorem II.1.11)

$$(\mathbf{1} + X)^{-1} = \sum_{k=0}^{\infty} (-1)^k X^{\otimes k}, \tag{8.5}$$

we obtain the desired inverse $a_0^{-1}(\mathbf{1} + b)^{-1}$ if the series converges. For this we observe that $\mathbf{1} - a_0^{-1}a$ has no part which is homogeneous of degree 0. Taking products of this element with itself, we only obtain contributions by homogeneous parts of higher degrees. Hence, if $N < \infty$ and we truncate, the series is just a polynomial. For $N = \infty$ we see similarly that after projecting to the factors $(\mathbb{R}^d)^{\otimes k}$ we again obtain only a polynomial. Hence also in this case the series converges as it (trivially) converges in every factor. Thus $a_0^{-1}a$ is invertible and the inverse depends continuously on a. Moreover, the set of units $\mathcal{T}^\infty(\mathbb{R}^d)^\times = (\pi_0^\infty)^{-1}(\mathbb{R} \setminus \{0\})$ is open. Summing up, this proves that $\mathcal{T}^\infty(\mathbb{R}^d)$ is a CIA. □

8.3 Corollary *For every $N \in \mathbb{N} \cup \{\infty\}$ and $d \in \mathbb{N}$, the group $(\mathcal{T}^N(\mathbb{R}^d))^\times = \{v \in \mathcal{T}^N(\mathbb{R}^d) \mid \pi_0^N(v) \neq 0\}$ is a regular locally exponential Lie group whose Lie algebra is $\mathcal{T}^N(\mathbb{R}^d)$ with the commutator bracket. Moreover, its Lie group exponential is given by*

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$\exp_\otimes : \mathcal{T}^N(\mathbb{R}^d) \to (\mathcal{T}^N(\mathbb{R}^d))^\times, \quad \exp_\otimes(v) = \sum_{k=0}^{\infty} \frac{v^{\otimes k}}{k!},$$

where $v^{\otimes k}$ is the kth power of v with respect to the algebra product.

Proof The Lie group structure was established in Example 3.4 and the Lie algebra computed in Exercise 3.2.8. Note that thanks to Lemma 8.2 the (truncated) tensor algebra is complete, whence its unit group is regular by Example 3.35. There we also mentioned that the evolution map Evol is given by the Volterra series (3.6). Plugging a constant path $t \mapsto v$ into the Volterra series we immediately obtain that $\exp_\otimes$ is the Lie group exponential (check this!). If $N < \infty$ the CIA is finite dimensional, and so $(\mathcal{T}^N(\mathbb{R}^d))^\times$ is a finite-dimensional Lie group and thus locally exponential. The case $N = \infty$ is much more involved and we refer to Glöckner and Neeb (2012, Theorem 5.6) for details. Note, however, that we will establish the convergence of the Lie group exponential and its inverse on a certain closed subspace in Lemma 8.5. □

Returning briefly to iterated integrals of a smooth path $X\colon [0,1] \to \mathbb{R}^d$, we shall now investigate iterated integrals of the path against itself. For this, identify $\mathbb{R}^d$ with the homogeneous elements of degree 1. Thus X becomes a smooth curve to the tensor algebra. However, for reasons which will become apparent in a moment, we are more interested in the smooth curve $DX(t) := (0, X'(t), 0, 0, \ldots) \in \mathcal{T}^N(\mathbb{R}^d)$. Applying the evolution to DX (viewed as a Lie algebra valued path) and sorting the result of the Volterra series by degree, we obtain

$$\mathrm{Evol}(DX)(t) = \left(t \mapsto \left(1, X_{0,t}, \int_0^t \int_0^r \mathrm{d}X \otimes \mathrm{d}X_r, \int_0^t \int_0^{r_2} \int_0^{r_1} \mathrm{d}X \otimes \mathrm{d}X \otimes \mathrm{d}X, \ldots\right)\right),$$

where we either truncate at $N < \infty$ or take the full tensor series for $N = \infty$. We have now used the evolution of the Lie groups to define a well-known object in the theory of rough paths.

8.4 Definition (Signature of a smooth path) Let $X\colon [0,1] \to \mathbb{R}^d$ be a smooth path. Then we define the *N-step signature*

$$S_N(X)_{s,t} := \left(1, X_{s,t}, \int_s^t \int_s^{r_1} \mathrm{d}X \otimes \mathrm{d}X, \int_s^t \int_s^{r_2} \int_s^{r_1} \mathrm{d}X \otimes \mathrm{d}X \otimes \mathrm{d}X, \ldots\right). \tag{8.6}$$

If $N = \infty$ we also write the shorter $S(X) := S_\infty(X)$ and say that $S(X)$ is the *signature of the smooth path X*.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

As $S_N(X)$ is the evolution of a smooth path it satisfies for every $N \in \mathbb{N} \cup \{\infty\}$ the Lie type differential equation

$$\begin{cases} \frac{d}{dt} S_N(X)_{0,t} = S_N(X)_{0,t} \otimes (dX)_t \quad t \in [0,1], \\ S_N(X)_{0,0} = \mathbf{1}. \end{cases} \tag{8.7}$$

Before we continue, it is important to stress that the signature can be defined not only for a smooth path. We shall see later (§8.3) that the signature exists for arbitrary rough paths. The signature of a path has turned out to be an immensely important object in the theory of rough paths and its applications. For example, the signature can be computed in advance for paths of interest and can then be applied in numerical analysis or for machine learning purposes. We refer to Chevyrev and Kormilitzin (2016) for an introduction.

We will generalise (8.7) in Exercise 8.2.4 and show that $S_N(X)_{s,t}$ can be recovered directly from $\mathrm{Evol}(DX)(t) = S_N(X)_{0,t}$ by virtue of Chen's relation $S_N(X)_{s,t} = S_N(X)_{s,u} \otimes S_N(X)_{u,t}$ for all $s < u < t$. However, our investigation so far also hints at the fact that the unit group of the tensor algebra is much larger than needed and contains many elements which will not turn out to be signatures of smooth paths. For example, the second level of the signature is a matrix (viewed as an element in $\mathbb{R}^d \otimes \mathbb{R}^d$). Its symmetric part is fixed by the shuffle of the level 1 part of the signature with itself (Chen's relation, cf, §8.4 for more on the shuffle product). Hence the second level uniquely is determined by its antisymmetric part which is in the stochastical theory interpreted as the Levy area. To capture these non-linear constraints, one restricts to a certain Lie subgroup of the unit group which expresses the geometric features of the signature. For this let us first study the restriction of the Lie group exponential.

8.5 Lemma *Let $N \in \mathbb{N} \cup \{\infty\}$ and $d \in \mathbb{N}$ and set $\mathcal{I}_N := (\pi_0^N)^{-1}(0)$. Then $\mathcal{I}_N$ is a Lie algebra ideal in $\mathcal{T}^N(\mathbb{R}^d)$ and the following maps are mutually inverse smooth diffeomorphisms*

$$\exp_N : \mathcal{I}_N \to \mathbf{1} + \mathcal{I}_N, \quad v \mapsto \sum_{0 \le n \le N} \frac{X^{\otimes n}}{n!},$$

$$\log_N : \mathbf{1} + \mathcal{I}_N \to \mathcal{I}_N, \quad \mathbf{1} + v \mapsto \sum_{0 \le n \le N} (-1)^{n+1} \frac{Y^{\otimes n}}{n}.$$

Proof Since the Lie bracket is given by the commutator, it is clear that $\mathcal{I}_N = (\pi_0^N)^{-1}(0)$ is a Lie ideal (i.e. $[v,w] \in \mathcal{I}_N$ if either v or w is in $\mathcal{I}_N$). First, note that since $v \in \mathcal{I}_N$ we have $\pi_0^N(v) = 0$ and thus $v^{\otimes k}$ does not contain contributions by homogeneous elements of degree less than k. In particular, we see that for every degree, the series $\exp_N$ and $\log_N$ reduce to polynomials in v. Thus both mappings are well-defined smooth mappings to $\mathcal{T}^N(\mathbb{R}^d)$. Inserting

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

and rearranging the formal power series into each other shows that $\exp_N$ and $\log_N$ are mutually inverse, hence diffeomorphisms. □

Note that $\mathbf{1}+\mathcal{I}_N$ is a subgroup of $\mathcal{T}^N(\mathbb{R}^d)^\times$ and a closed submanifold (as a closed affine subspace of $\mathcal{T}^N(\mathbb{R}^d)$). Thus, in particular, it is a closed Lie subgroup of the unit group. Since the exponential series yields the Lie group exponential of the unit group, we can interpret Lemma 8.5 as the statement that the Lie group exponential of $\mathbf{1}+\mathcal{I}_N$ is a diffeomorphism from the Lie algebra onto the group. However, the Lie group we are after is yet a smaller Lie subgroup of $\mathbf{1}+\mathcal{I}_N$ which nevertheless contains all signatures.

8.6 Definition Let $N \in \mathbb{N}\cup\{\infty\}, d \in \mathbb{N}$. Then we define $\mathfrak{g}^N(\mathbb{R}^d)$ as the smallest closed Lie subalgebra generated by the homogeneous elements of degree 1 in $\mathcal{T}^N(\mathbb{R}^d)$. Explicitly, these algebras are constructed as follows: Set $\mathcal{P}^1(\mathbb{R}^d) := \mathbb{R}^d \subseteq \mathcal{T}^N(\mathbb{R}^d)$ (via the canonical identification). Then define recursively:

$$\mathcal{P}^{n+1}(\mathbb{R}^d) := \mathcal{P}^n(\mathbb{R}^d) + \operatorname{span}\{[x,y] \mid x \in \mathbb{R}^d, y \in \mathcal{P}^n(\mathbb{R}^d)\},$$
$$\mathcal{P}^\infty(\mathbb{R}^d) := \{(0,P_1,P_2,\ldots) \mid \text{for all } i \in \mathbb{N}, P_i \in (\mathbb{R}^d)^{\otimes i} \cap \mathcal{P}^n(\mathbb{R}^d) \text{ for some } n \in \mathbb{N}\}.$$

Elements in $\mathcal{P}^n(\mathbb{R}^d)$ are called *Lie polynomials* while elements in $\mathcal{P}^\infty(\mathbb{R}^d)$ are called *Lie series*.[4] Then we set $\mathfrak{g}^N(\mathbb{R}^d) := \overline{\mathcal{P}^N(\mathbb{R}^d)}$ (where the bar denotes topological closure) and observe that these spaces are closed Lie subalgebras of $\mathcal{I}_N$ for $N \in \mathbb{N}\cup\{\infty\}$.

If we take now the image of $\mathfrak{g}^N(\mathbb{R}^d)$ under the exponential map $\exp_N$ we obtain a closed subset $G^N(\mathbb{R}^d) := \exp_N(\mathfrak{g}^N(\mathbb{R}^d))$ of the unit group. It is non-trivial to see that $G^N(\mathbb{R}^d)$ forms a group under tensor multiplication. The classical proof for this fact employs the Baker–Campbell–Hausdorff series as an essential tool. As this would lead us too far from our objects of interest, we will import this result and investigate just its differentiability.

8.7 Proposition *The set $G^N(\mathbb{R}^d)$ is a closed Lie subgroup of the unit group $(\mathcal{T}^N(\mathbb{R}^d))^\times$ for all $N \in \mathbb{N}\cup\{\infty\}$. Moreover, this structure turns it into a (locally) exponential Lie group.*

Proof We have seen already that $G^N(\mathbb{R}^d)$ is a closed subset. It is a subgroup by Reutenauer (1993, Corollary 3.3). To see that it is a Lie group, we have to show that it is a submanifold. Exploit that $\mathcal{T}^N(\mathbb{R}^d)$ is locally exponential (see

[4] The name hails from the custom of writing a Lie series in the form of a formal power series instead of the sequence representation we have chosen. We refer to Reutenauer (1993, Chapter 1) for more information.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Corollary 8.3). Hence there is a 0-neighbourhood $V \subseteq \mathcal{T}^N(\mathbb{R}^d)$ such that $\exp_\otimes$ restricts to a diffeomorphism on V. By construction we have $\exp_\otimes(\mathfrak{g}^N(\mathbb{R}^d)) = \exp_N(\mathfrak{g}^N(\mathbb{R}^d)) = G^N(\mathbb{R}^d)$, whence $\exp_\otimes(V \cap \mathfrak{g}^N(\mathbb{R}^d) = \exp_\otimes(V) \cap G^N(\mathbb{R}^d)$ yields a submanifold chart (φ, V) for $G^N(\mathbb{R}^d)$ around the identity. Exploiting the idea that $\mathcal{T}^N(\mathbb{R}^d)^\times$ is a Lie group, we see that a submanifold atlas for $G^N(\mathbb{R}^d)$ is then given by $\varphi_g : V \to G^N(\mathbb{R}^d)$, $x \mapsto g\varphi(x)$, $g \in G^N(\mathbb{R}^d)$. We conclude that $G^N(\mathbb{R}^d)$ is a closed Lie subgroup of the unit group with Lie algebra $\mathfrak{g}^N(\mathbb{R}^d)$. To see that it is locally exponential, it suffices to notice that $\exp_\otimes$ restricts to the Lie group exponential of $G^N(\mathbb{R}^d)$. This is due to the fact that the Lie group exponential is defined via solution to certain differential equations. Given initial values in the closed subspace $\mathfrak{g}^N(\mathbb{R}^d)$, the solutions of the equation in the unit group already stay in $G^N(\mathbb{R}^d)$, whence they solve the differential equation in the subgroup. □

The algebraic arguments in this section were just cited from the literature as we wished to keep the exposition simple. However, the algebraic structure is the key to understanding the Lie groups at hand, since, as a consequence of Lemma 8.5, the Lie groups are globally diffeomorphic to their Lie algebra via the exponential map. For the next result, we assume familiarity with projective limits (see e.g. Hofmann and Morris, 2007, Chapter 1).

8.8 Proposition *Let $m, n \in \mathbb{N} \cup \{\infty\}$ such that $m \geq n$.*

(a) *Then the canonical projection $\pi_n^m : \mathcal{T}^m(\mathbb{R}^d) \to \mathcal{T}^n(\mathbb{R}^d)$ is a morphism of locally convex algebras, which restricts to a Lie group morphism $p_n^m : G^m(\mathbb{R}^d) \to G^n(\mathbb{R}^d)$.*

(b) *We obtain a commutative diagram of Lie groups and their associated Lie algebras*

$$\begin{array}{ccccccccc}
\{0\} \times \mathbb{R}^d = \mathfrak{g}^1(\mathbb{R}^d) & \xleftarrow[\mathbf{L}(p_1^2)]{} & \mathfrak{g}^2(\mathbb{R}^d) & \xleftarrow[\mathbf{L}(p_2^3)]{} & \mathfrak{g}^3(\mathbb{R}^d) & \longleftarrow & \cdots \\
\downarrow{\scriptstyle \exp_1} & & \downarrow{\scriptstyle \exp_2} & & \downarrow{\scriptstyle \exp_3} & & \\
\{1\} \times \mathbb{R}^d = G^1(\mathbb{R}^d) & \xleftarrow[p_1^2]{} & G^2(\mathbb{R}^d) & \xleftarrow[p_2^3]{} & G^3(\mathbb{R}^d) & \longleftarrow & \cdots,
\end{array} \tag{8.8}$$

where the upper row is a projective system of locally convex Lie algebras whose limit is $\mathfrak{g}^\infty(\mathbb{R}^d)$.

Proof From the definition of the product in the tensor algebra it is clear that the π_n^m are algebra morphisms. As π_n^m is just the projection from a product onto some of its components, it is continuous in the product topology, hence a morphism of locally convex algebras. Note that this entails that π_n^m restricts to a morphism of locally convex Lie algebras $q_n^m : \mathfrak{g}^m(\mathbb{R}^d) \to \mathfrak{g}^n(\mathbb{R}^d)$ and a

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

morphism of Lie groups $p_n^m : G^m(\mathbb{R}^d) \to G^n(\mathbb{R}^d)$ (it is smooth as the restriction of a continuous linear map to closed submanifolds). Since π_n^m is linear, we have $\mathbf{L}(p_m^n) = T_{\mathbf{1}}\pi_n^m|_{\mathfrak{g}^\infty(\mathbb{R}^d)} = q_n^m$. Thus the naturality of the exponential map (3.8) yields $p_n^m \circ \exp_m = \exp_n \circ \mathbf{L}(p_m^n)$. Summing up, this proves (a) and establishes the commutativity of (8.8).

For part (b) let us note that (a) establishes that the upper row of (8.8) is a projective system of locally convex Lie algebras (as $\mathbf{L}(p_i^j) \circ \mathbf{L}(p_k^i) = \mathbf{L}(p_k^j)$ and these mappings are continuous morphisms of Lie algebras). Consider now $x \in \mathfrak{g}^\infty(\mathbb{R}^d)$. Its projection $\pi_n^\infty(x)$ is contained in $\mathfrak{g}^n(\mathbb{R}^d)$ (this is clear for a Lie series and follows by considering converging sequences for the elements in the closure since the Lie algebras $\mathfrak{g}^n(\mathbb{R}^d)$ are closed). Since π_n^∞ restricts to the Lie algebra morphism $\mathbf{L}(p_n^\infty)$ on $\mathfrak{g}^\infty(\mathbb{R}^d)$, this implies that $\mathfrak{g}^\infty(\mathbb{R}^d)$ is the projective limit of the projective system in the category of Lie algebras. In addition, the product topology on $\mathcal{T}^\infty(\mathbb{R}^d)$ is the projective limit of the locally convex spaces $\mathcal{T}^n(\mathbb{R}^d)$, whence the topology on $\mathfrak{g}^\infty(\mathbb{R}^d)$ is the locally convex projective limit topology induced by the projective system. In conclusion, $\mathfrak{g}^\infty(\mathbb{R}^d)$ is the projective limit in the category of locally convex Lie algebra. □

8.9 Remark (Projective limits of finite-dimensional Lie groups) In Proposition 8.8 we exploited that the projective limit of locally convex Lie algebras can be described as the projective limit of Lie algebras with the (locally convex) projective limit topology. Thanks to the commutativity of (8.8), the lower row also forms a projective system of finite-dimensional Lie groups. Projective limits for Lie groups may not exist (while they always exist in the category of topological groups). Topological groups which are projective limits of finite-dimensional Lie groups are called *pro-Lie groups* (Hofmann and Morris, 2007). Hence (8.8) shows that $G^\infty(\mathbb{R}^d)$ is a pro-Lie group which is simultaneously a Lie group. This situation has been studied in Hofmann and Neeb (2009). It is worth mentioning that groups with these properties inherit a surprising amount of structure from the finite-dimensional Lie groups which were used in their construction.

Summing up, the truncated groups are closely connected to their projective limit $G^\infty(\mathbb{R}^d)$. Moreover, (truncated) signatures of smooth paths extend naturally to the projective limit. In the next section we will review the concept of a rough path, which, in a certain sense, generalises the signature for paths of low regularity. For this, the geometry of the groups $G^N(\mathbb{R}^d)$ will be instrumental.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Exercises

8.2.1 Declare a Hilbert space structure on $\mathcal{T}^N(\mathbb{R}^d)$ for $N \in \mathbb{N}$ by defining the canonical basis elements $e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_k}$ to be orthonormal. Show that:

(a) The norm corresponding to the inner product satisfies $\|v \otimes w\| \leq \|v\| \cdot \|w\|$ and $\|v \otimes w\| = \|w \otimes v\|$. Deduce that $\mathcal{T}^N(\mathbb{R}^d)$ becomes a Banach algebra.

(b) Identifying the homogeneous elements of degree k with elements $\mathbb{R}^{d^k}$ by sending the canonical bases to each other, the resulting isomorphism is an isometry of Hilbert spaces.

8.2.2 Consider the Step 2 truncated tensor algebra $\mathcal{T}^2(\mathbb{R}^d) = \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \otimes \mathbb{R}^d$.

(a) Show that the multiplication of the truncated tensor algebra is given by

$$(a,b,c) \cdot (x,y,z) = (ax, ay + xb, az + xc + b \otimes y)$$
$$\text{and} \quad (1,b,c)^{-1} = (1,-b,-c + b \otimes b).$$

Deduce that if $d \neq 1$, the product is not commutative.

(b) Let $X\colon [0,1] \to \mathbb{R}^d$ be a smooth path and $\mathbb{X}_{s,t} := \int_s^t X_{s,r} \otimes X'_r \mathrm{d}r$. Define $\mathbf{X}_{s,t} := (1, X_{s,t}, \mathbb{X}_{s,t}) \in \mathcal{T}^2(\mathbb{R}^d)$. Establish that X satisfies Chen's relation (8.4)

$$\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u} \otimes X_{u,t}, \quad \text{for all } u \in [s,t].$$

Prove then that $\mathbf{X}_{s,t} = \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t}$ (also called Chen's relation).

8.2.3 Show that for $N \in \mathbb{N}$ the algebra $\mathcal{T}^N(\mathbb{R}^d)$ is a quotient of $\mathcal{T}^\infty(\mathbb{R}^d)$ modulo the algebra ideal $\mathcal{I}_N = \{(x_k)_{k\in\mathbb{N}} \in \mathcal{T}^\infty(\mathbb{R}^d) \mid x_1 = \cdots = x_N = 0\}$.

8.2.4 Let $X\colon [0,1] \to \mathbb{R}^d$ be a smooth path, $N \in \mathbb{N} \cup \{\infty\}$ and $DX\colon [0,1] \to \mathcal{T}^N(\mathbb{R}^d)$, $t \mapsto (0, X'_t, 0, \ldots)$. Show that

(a) for fixed s the signature satisfies the differential equation

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} S_N(X)_{s,t} = S_N(X)_{s,t} \otimes (DX)_t, & s < t \leq 1, \\ S_N(X)_{s,s} = \mathbf{1} = (1,0,\ldots), \end{cases}$$

(b) the signature satisfies Chen's relation

$$S_N(X)_{s,t} = S_N(X)_{s,u} \otimes S_N(X)_{u,t}, \qquad 0 \leq s \leq u \leq t \leq 1.$$

Hint: It suffices to prove this for every projection of $S_N(X)$ to $(\mathbb{R}^d)^{\otimes k}$ and consider the iterated integral of $\mathrm{d}X_{r_1} \otimes \cdots \otimes \mathrm{d}X_{r_k}$ over the simplex $\Delta^N = \{s < r_1 < r_2 < \cdots < r_n < t\}$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(c) Deduce that $S_N(X)_{s,t} = S_N(X)_{0,s}^{-1} \otimes S_N(X)_{0,t}$, whence the signature can be recovered from the curve $\mathrm{Evol}(DX)$ via the group operations.

8.2.5 Let $a \in \mathcal{T}^N(\mathbb{R}^d)$, $N \in \mathbb{N} \cup \{\infty\}$ such that $a_0 := \pi^N(a) \neq 0$. Write $a = a_0(\mathbf{1} + b)$ and prove that the Neumann inverse (8.5) yields $a^{-1} = a_0^{-1}(\mathbf{1} + b)^{-1} = a_0^{-1} \sum_{k=1}^{\infty} (-1)^k b^{\otimes k}$.

8.3 A Rough Introduction to Rough Paths

In this section we will recall the notion of a rough path. The main idea of rough path theory is that paths of much lower regularity than being smooth can be augmented with extra information replacing the iterated integrals we studied in the last section. Indeed the basic idea is to declare the signature to be the object of interest and define signature-like objects in the tensor algebra. While we will present the basic theory of rough paths, we recommend one of the excellent introductions to rough path theory available (see e.g. Friz and Victoir, 2010; Friz and Hairer, 2020) for more in-depth information.

As a starting point, let us formalise the properties observed for the signature in the last section.

8.10 Definition Let $\Delta := \{(s,t) \in [0,1] \mid 0 \leq s \leq t \leq 1\}$ be the *standard simplex*, $N \in \mathbb{N} \cup \{\infty\}$ and $d \in \mathbb{N}$. We call a map

$$\mathbf{X} \colon \Delta \to \mathcal{T}^N(\mathbb{R}^d), \quad \mathbf{X}_{s,t} = (X_{s,t}^0, X_{s,t}^1, X_{s,t}^2, \ldots)$$

multiplicative functional[5] of degree N if $X_{s,t}^0 = 1$ for all $(s,t) \in \Delta$ and the map satisfies

$$\textit{Chen's relation}: \quad \mathbf{X}_{s,t} = \mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} \text{ for all } s,u,t \in [0,1], s \leq u \leq t. \quad (8.9)$$

Note that Chen's relation for the first two (non-trivial) components of $\mathbf{X}$ reduces to (cf. Exercise 8.2.2)

$$X_{s,t}^1 = X_{s,u}^1 + X_{u,t}^1, \qquad X_{s,t}^2 = X_{s,u}^2 + X_{u,t}^2 + X_{s,u}^1 \otimes X_{u,t}^1. \quad (8.10)$$

Moreover, Lemma 8.2 (a) shows that every multiplicative functional is invertible in the tensor algebra, whence Chen's relation entails $\mathbf{X}_{s,t} = \mathbf{X}_{0,s}^{-1} \otimes \mathbf{X}_{0,t}$. So instead of a multiplicative functional, we may think of the path $[0,1] \to \mathcal{T}^N(\mathbb{R}^d)$, $t \mapsto \mathbf{X}_{0,t}$. Hence the term 'rough path' will make more sense once

[5] There is a deeper story going on which motivates the term 'multiplicative functional'. We will briefly discuss this in §8.4.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

we define it. For the first level, however, this translates to the statement that there is a path $X\colon [0,1] \to \mathbb{R}^d$ with $X^1_{s,t} = X_t - X_s$.

8.11 Example In the previous section we saw that the signature of a smooth path yields a multiplicative functional. For example, we can augment the zero-path in different ways to generate multiplicative functionals: $\mathbf{X}_{s,t} := (1, 0, (t-s)w)$ is a multiplicative functional of degree 2 for any $w \in \mathbb{R}^d \otimes \mathbb{R}^d$. Note that it coincides with the level 2-signature of the zero path only if w is the zero element. More generally, for any function $F\colon [0,1] \to (\mathbb{R}^d)^{\otimes N}$ the map $\mathbf{X}_{s,t} = (1, 0, \dots, 0, F(t) - F(x)) \in \mathcal{T}^N(\mathbb{R}^d)$ is a multiplicative functional (see Exercise 8.3.1).

Now we need to add an analytic condition to the algebraic objects we just defined. In this case we wish to consider rough signals in the framework of Hölder regular paths.

8.12 Definition Let $X\colon [0,1] \to E$ be a continuous map with values in a Banach space $(E, \|\cdot\|)$. We say X is an α-Hölder path for $0 < \alpha < 1$ if

$$\|X\|_\alpha := \sup_{\substack{t,s\in[0,1],\\ t\neq s}} \frac{\|X_t - X_s\|}{|t-s|^\alpha} < \infty.$$

Denote by $C^\alpha([0,1], E)$ *the space of all α-Hölder continuous functions*.

It is well known (see Exercise 8.3.2) that $C^\alpha([0,1], E)$ is a Banach space on which $\|\cdot\|_\alpha$ is a continuous seminorm (it annihilates all constant paths).

8.13 Remark Again we will be restricted here to paths on the interval $[0,1]$. The results carry over to any interval by composing the Hölder paths with a reparametrisation of the interval. Note, however, that the Hölder norm is *not* invariant under reparametrisation. This is one reason why in the rough paths literature one often considers the (more or less) equivalent formulation via p-variation paths for $p > 1$. The p-variation norm is invariant under reparametrisation; see Friz and Victoir (2010).

Due to Young's theorem, it is not possible to augment an α-Hölder path with iterated integrals against itself if $\alpha < 1/2$. The core idea of rough path theory is to augment an α-Hölder path with additional information to make integration against it feasible and circumvent the restrictions of Young integration theory. To this end we augment the concept of a multiplicative functional with a Hölder condition.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

8.14 Definition Fix $\alpha \in]0,1[$ and $N \geq \lfloor 1/\alpha \rfloor$. An $\mathbb{R}^d$-valued *α-rough path* consists of a multiplicative functional $\mathbf{X} := (1, X^1, X^2, X^3, \dots, X^N)$ such that the kth component X^k is 'k-times α Hölder continuous', that is,

$$\|X^k_{s,t}\| \lesssim |t-s|^{k\alpha}, \text{ where '}\lesssim\text{' means inequality up to a constant.} \tag{8.11}$$

If, in addition, the α-rough path takes values in $G^N(\mathbb{R}^d)$, we call $\mathbf{X}$ a *weakly geometric rough path*. We define $\mathrm{RP}^\alpha([0,1],\mathbb{R}^d)$ as the *set of weakly geometric α-rough paths* (we shall see in Theorem 8.15 that it makes sense to drop N from the notation for $\mathrm{RP}^\alpha([0,1],\mathbb{R}^d)$).

The higher levels of a rough path are not given by increments of a function, whence they do not become constant even if k-times α-Hölder means that the Hölder index satisfies $k\alpha > 1$. Lyons' original concept of rough path does not require the rough path to take its image in the group $G^N(\mathbb{R}^d)$. However, it has turned out that the general notion is not sufficient to solve non-linear rough differential equations (see e.g. Friz and Victoir, 2010; Friz and Hairer, 2020). For this purpose one should consider weakly geometric rough paths and we will do this in the rest of the section. To ease notation we shall (unless we explicitly say otherwise) only consider weakly geometric rough paths and simply call them rough paths. Let us recall that the cutoff level N in the definition of a rough path is unimportant as long as $N \geq \lfloor 1/\alpha \rfloor$.

8.15 Theorem (Lyons' lifting theorem (Friz and Victoir, 2010, Theorem 9.5)) *Let $\alpha \in]0,1[$ and $\lfloor 1/\alpha \rfloor \leq n$. Assume that $\mathbf{X}\colon \Delta \to G^n(\mathbb{R}^d)$ is an α-rough path. Then there exists a unique α-rough path $\mathbf{X}^{n+1}\colon \Delta \to G^{n+1}(\mathbb{R}^d)$ extending $\mathbf{X}$, that is, if $\pi^{n+1}_n\colon \mathcal{T}^{n+1}(\mathbb{R}^d) \to \mathcal{T}^n(\mathbb{R}^d)$ is the canonical projection, then $\pi^{n+1}_n(\mathbf{X}^{n+1}) = \mathbf{X}$. A rough path extending $\mathbf{X}$ in this way is called a* Lyons lift *of $\mathbf{X}$.*

Thus every α-rough path with values in $G^n(\mathbb{R}^d)$ extends to an α-rough path with values in $G^m(\mathbb{R}^d), m \geq n \geq \lfloor 1/\alpha \rfloor$.

As a consequence of Theorem 8.15, every α-rough path can be obtained as a restriction of an α-rough path with values in $G^\infty(\mathbb{R}^d)$, or conversely as an extension of an α-rough path with values in $G^{\lfloor 1/\alpha \rfloor}(\mathbb{R}^d)$. The remarkable fact here is of course that the extension is uniquely determined once information up to level $\lfloor 1/\alpha \rfloor$ is available.

8.16 Example We have seen that for a smooth path X and $N \in \mathbb{N} \cup \{\infty\}$, the signature $S_N(X)$ is a multiplicative functional. More generally, if we start with an α-Hölder path for $\alpha > 1/2$, we can compute its signature using iterated Young integrals. It is then a consequence of Young's inequality (Young, 1936)

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

that the resulting multiplicative functional is indeed an α-rough path (see Friz and Hairer, 2020, Section 4, for a detailed discussion).

For paths of lower regularity, the idea is that the components X^k of a rough path $\mathbf{X}$ replace the (in general, ill-defined) iterated integrals of the path giving the first level increments. Due to its importance, let us stress it here again explicitly: If $\mathbf{X} = (1, X^1_{s,t}, X^2_{s,t})$ is an α-rough path for $\alpha \in]1/3, 1/2[$ and $X^1_{s,t} = X_t - X_s$, one should interpret the second level as

$$\int_s^t X_{u,s} \mathrm{d}X_u := X^2_{s,t}, \qquad 0 \le s \le t \le 1,$$

where the left-hand side is defined via the right-hand and not the other way around! The levels of an α-rough path thus encode information similar to the iterated integrals in the signature of a smooth path.

It is not obvious at all whether a given α-Hölder path with values in $\mathbb{R}^d$ can be enhanced to yield an α-rough path. Thanks to a result by Lyons and Victoire, this can be achieved in many cases through an abstract extension result. If $1/\alpha \notin \mathbb{N}$, every α-Hölder path can be enhanced (non-uniquely) to an α-Hölder rough path (this is the Lyons–Victoire lifting theorem; Lyons and Victoir, 2007, Theorem 1). As an example of the non-uniqueness of the extension let us mention the following example for rough paths arising from Brownian motion.

8.17 Example (Brownian motion as a rough path) Brownian motion models the (random) movement of a particle in a fluid. It can be modelled as a stochastic process $B\colon \Omega \times [0,1] \to \mathbb{R}^d$, with independent Gaussian increments and continuous sample paths. Here $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. However, we refer to the stochastic literature for explanations and more details. One can show that for partitions P of $[0,1]$, the Riemann sum

$$\int_0^1 B_r \otimes d^\theta B_r = \lim_{|P|\to 0} \sum_{P=(t_i)} B_{t_i+\theta(t_{i+1}-t_i)} \otimes B_{t_i,t_{i+1}}$$

converges in $L^2(\Omega, \mathbb{R}^{d\times d})$ for mesh size $|P|$ converging to 0 and $\theta \in [0,1]$. Contrary to usual Riemann–Lebesgue integration theory, the integral $\int B_r \otimes d^\theta B_r$ is *not* independent of θ. The choice $\theta = 0$ leaves the martingale structure invariant and is referred to as the Itô integral, whereas $\theta = \frac{1}{2}$ is compatible with regular calculus rules and is referred to as the Stratonovich integral. It is well known that B is an α-Hölder path for every $\alpha \in]1/3, 1/2[$ and if we write $B^{\mathrm{Itô}}$ and B^{Strat} for the Itô and the Stratonovich integrals, we obtain two α-rough paths $(1, B, B^{\mathrm{Itô}})$ and $(1, B, B^{\mathrm{Strat}})$. Note that the rough path obtained via the Itô integral does not take its values in $G^2(\mathbb{R}^d)$ (i.e. it is not a weakly geometric rough path), while the one obtained from Stratonovich

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

integration takes values in $G^2(\mathbb{R}^d)$ and is weakly geometric. We refer to Friz and Hairer (2020, Chapter 3) for a detailed discussion of these examples.

The reader may wonder now in what sense a rough path is a path, that is, can we interpret it as a Hölder continuous path $[0,1] \to G^N(\mathbb{R}^d)$? For this we have to consider the metric geometry of the groups $G^N(\mathbb{R}^d)$ for $N < \infty$. However, as a first step to make sense of the following constructions, we need the following result (whose full proof we postpone to Remark 8.25).

8.18 Lemma *Let* $\mathbf{X}\colon \Delta \to G^N(\mathbb{R}^d), \mathbf{X} = (1, X^1, X^2, \dots)$ *be an* α*-rough path. Taking the pointwise inverse, the map* $\mathbf{X}^{-1}\colon \Delta \to G^N(\mathbb{R}^d)$, $t \mapsto (\mathbf{X}_t)^{-1} = (1, X_t^{-1}, X_t^{-2}, \dots)$ *is graded* $k\alpha$*-Hölder, that is,* $\|X^{-k}\|_{k\alpha} < \infty$ *for all* $k \in \mathbb{N}$.

Proof We will only prove the case where $N = 2$ and $d = 2$. Without more techniques from §8.4 the general case turns out to be quite involved. Note that for $N = 2$, thanks to Exercise 8.2.2(a), we have $\mathbf{X}^{-1} = (1, -X^1, -X^2 + X^1 \otimes X^1)$. Thus the degree 1 component is α-Hölder since X^1 is α-Hölder. Immediately, we see that the degree 2 component might not be 2α-Hölder, as $X^1 \otimes X^1$ is a product of α-Hölder functions. To circumvent this, we express $\mathbf{X}$ with the help of the standard basis e_1, e_2 of $\mathbb{R}^2$ as $\mathbf{X} = (1, x_1 e_1 + x_2 e_2, \sum_{1 \le i,j \le 2} y_{ij} e_i \otimes e_j)$. Working out the constraints imposed by $G^2(\mathbb{R}^2) = \exp_2(\mathfrak{g}^2(\mathbb{R}^2))$ (the reader should check this!), we find that

$$x_1 x_2 = y_{12} + y_{21}, \quad x_1^2 = 2y_{11}, \quad x_2^2 = 2y_{22}.$$

Plugging this into the inversion formula we find for the degree 2 component,

$$\begin{aligned} X^{-2} &= -\sum_{1 \le i,j \le 2} y_{ij} e_i \otimes e_j + x_1^2 e_1 \otimes e_1 + x_1 x_2 e_1 \otimes e_2 + x_1 x_2 e_2 \otimes e_1 + x_2^2 e_2 \otimes e_2 \\ &= y_{11} e_1 \otimes e_1 + y_{22} e_2 \otimes e_2 + y_{12} e_2 \otimes e_1 + y_{21} e_1 \otimes e_2. \end{aligned}$$

Thus the second component is 2α-Hölder as it contains only contributions from the 2α-Hölder maps y_{ij} comprising the second level of $\mathbf{X}$. This happens in general, as the non-linear constraints imposed by $G^N(\mathbb{R}^d)$ allow us to rewrite the kth degree component of $\mathbf{X}^{-1}$ as a sum of the functions comprising the kth degree component of $\mathbf{X}$ (whence they preserve the Hölder condition). □

The discussion of the $N = 2, d = 2$ case reveals that Lemma 8.18 will, in general, be false if the rough path is not weakly geometric. We now define a suitable metric on the group $G^N(\mathbb{R}^d)$ for which every α-rough path will correspond to a Hölder function.

8.19 Definition Consider for $N \in \mathbb{N}$ the subset $\mathcal{T}_1^N(\mathbb{R}^d) = \{x \in \mathcal{T}^N(\mathbb{R}^d) \mid \pi_0^N(x) = 1\}$. Let $x = (1, x^1, x^2, \dots, x^N) \in \mathcal{T}_1^N(\mathbb{R}^d)$; then we define

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$|x| := \max_{k=1,\dots,N} \{(k!\|x^k\|)^{1/k}\} + \max_{k=1,\dots,N} \{(k!\|(x^{-1})^k\|)^{1/k}\}, \tag{8.12}$$

where $\|\cdot\|$ is the norm we have chosen on $\mathcal{T}^N(\mathbb{R}^d)$. Then we set $\rho_N(x,y) := |x^{-1} \otimes y|$. Identifying $\mathcal{T}_1^N(\mathbb{R}^d)$ with the vector space $\mathcal{I}_N$ (by subtracting $\mathbf{1}$), it is easy to see that $|\cdot|$ becomes a norm and ρ_N induces a left-invariant, symmetric and subadditive metric on $G^N(\mathbb{R}^d)$ (see Exercise 8.3.4 for the details).

The metric ρ_N turns the group $G^N(\mathbb{R}^d)$ into a homogeneous group in the sense of Folland and Stein (1982).[6] Now if $\mathbf{X}$ is an α-rough path (with values in $G^N(\mathbb{R}^d)$), then the path $x_t := \mathbf{X}_{0,t}$ satisfies

$$\|x\|_\alpha^{\rho_N} := \sup_{\substack{t \neq s \\ s,t \in [0,1]}} \frac{\rho_N(x_s, x_t)}{|t-s|^\alpha} = \sup_{\substack{t \neq s \\ s,t \in [0,1]}} \frac{\rho_N(\mathbf{1}, \mathbf{X}_{s,t})}{|t-s|^\alpha} < \infty. \tag{8.13}$$

In other words, an α-rough path is an α-Hölder continuous path with values in metric space $(G^N(\mathbb{R}^d), \rho_N)$. This statement hinges on the rough path taking values in $G^N(\mathbb{R}^d)$ (i.e. the path being weakly geometric) as we need Hölder continuity of the pointwise inverse $\mathbf{X}^{-1}$.

Indeed due to Chen's relation this is an equivalent point of view, up to forgetting the starting point of the path in $\mathbb{R}^d$. Unfortunately this point of view is limited to the truncated groups, as the metric (8.12) does not extend to $G^\infty(\mathbb{R}^d)$.

8.20 Remark In Remark 8.9 we saw that $G^\infty(\mathbb{R}^d)$ is the projective limit (as a Lie group) of the truncated groups $G^N(\mathbb{R}^d)$. Moreover, this Lie group is modelled on the Fréchet space $\mathfrak{g}^\infty(\mathbb{R}^d)$. So, in view of the characterisation of rough paths as Hölder continuous paths in the truncated groups, the question is of course whether we can use the metric on $\mathfrak{g}^\infty(\mathbb{R}^d)$ to construct a suitable metric dist which allows us to cast α-rough path as an α-Hölder path with values in the metric space $(G^\infty(\mathbb{R}^d), \text{dist})$.

This problem was considered in Le Donne and Züst (2021) starting from the *Carnot–Caratheodory metric* on $G^N(\mathbb{R}^d), N < \infty$. It is defined as $d_{\text{CC}}^N(y,z) := d_{\text{CC}}^N(\mathbf{1}, y^{-1} \otimes z)$ and

$$d_{\text{CC}}^N(\mathbf{1}, y) := \inf \left\{ \int_0^1 \|X_t'\| \, \mathrm{d}t \,\middle|\, X \in C([0,1], \mathbb{R}^d), \begin{array}{l} X_0 = 0, \ X_t \text{ has bounded variation} \\ \text{and } y = S_N(X)_{0,1} \end{array} \right\},$$

where the signature of the bounded variation path is defined as in (8.6) using (iterated) Riemann–Stieltjes integrals. We will abbreviate this metric as the

[6] The groups $G^N(\mathbb{R}^d)$ possess many strong structural properties which are exploited in rough path theory. They are homogeneous groups and connected and simply connected nilpotent Lie groups (i.e. Carnot groups, which are studied in sub-Riemannian geometry; see Le Donne and Züst (2021)). The details are beyond the scope of this chapter and we refer the reader to the literature.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

CC-metric and note that it is not straightforward to prove that it is a metric. Moreover, one can show that the CC-metric is equivalent to the metric ρ_N (see Friz and Victoir, 2010, 7.5.4 for details on the CC-metric). Now one can argue that the metric d_∞ should correspond to the metric on the projective limit of the system $(G^N(\mathbb{R}^d), \mathbf{1}, d_{CC}^N)$ in the category of (pointed) metric spaces (with morphisms given by submetries). However, one of the main results of Le Donne and Züst (2021) is the somewhat surprising insight that the limiting object $(G_\infty, \mathbf{1}, d_\infty)$ cannot be a topological group in the topology induced by d_∞ (let us note that $G_\infty \neq G^\infty(\mathbb{R}^d)$).

Hence there seems to be no straightforward way to construct a metric on the projective limit which simultaneously captures the desired convergence and geometry, is left-invariant and turns the projective limit into a Lie group.

Summing up, rough paths can naturally be identified as maps with values in the infinite-dimensional Lie group $G^\infty(\mathbb{R}^d)$. This group is closely connected to the truncated groups $G^N(\mathbb{R}^d)$ and inherits many properties from the projective system of these groups. However, there are important geometric properties connected to the sub-Riemannian geometry of the groups $G^N(\mathbb{R}^d)$, which have no counterpart in the infinite-dimensional group. This leaves us at an uncomfortable situation: If it is enough to work with the finite-dimensional groups, why bother with the more complicated situation of $G^\infty(\mathbb{R}^d)$?

One reason to care about the infinite-dimensional group is that it hosts all rough paths regardless of their regularity (recall that an α-rough path can only be defined on the group $G^N(\mathbb{R}^d)$, where $N \geq \lfloor 1/\alpha \rfloor$). It would be interesting to understand the geometry and manifold structure of the set of all elements which can be reached by α-rough paths (the construction of tangent spaces to α-rough paths for $\alpha \in]1/3, 1/2[$ in Qian and Tudor (2011) can be understood as a step in this direction). Another reason might be that one could be interested in different flavours of rough paths such as Gubinelli's branched rough paths (Gubinelli, 2010). The properties of branched rough paths require a different geometry and one has to replace the groups $G^N(\mathbb{R}^d)$. We shall describe the general construction in the next section (but mention that for the branched rough paths the construction actually yields isomorphic groups due to a deep algebraic result).

Exercises

8.3.1 Let $N \in \mathbb{N}$ and $\mathbf{X}, \mathbf{Y} \colon \Delta \to \mathcal{T}^N(\mathbb{R}^d)$ be multiplicative functionals which agree up to degree $N-1$ (i.e. their projections onto the components up to mth level are equal for all $m \leq N$). Then prove that

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(a) The difference function $F_{s,t} = X^N_{s,t} - Y^N_{s,t} \in (\mathbb{R}^d)^{\otimes N}$ is additive in the sense that for all $s \le u \le t$ one has $F_{s,t} = F_{s,u} + F_{u,t}$.

(b) If $F\colon \Delta \to (\mathbb{R}^d)^{\otimes N}$ is an additive function (in the above sense), then $(s,t) \mapsto \mathbf{X}_{s,t} + F_{s,t}$ is also an additive functional (where addition is addition in the tensor algebra).

8.3.2 We consider the α-Hölder space $C^\alpha([0,1],E)$ for $n \in \mathbb{N}$, $0 < \alpha < 1$ for $(E, \|\cdot\|)$ a Banach space. Show that the map

$$\|x\|_\alpha := \sup_{\substack{t,s\in[0,1],\\ t\neq s}} \frac{\|x_t - x_s\|}{|t-s|^\alpha}$$

is a semi-norm on $C^\alpha([0,1],E)$ which is not a norm. Then deduce that

$$\|x\|_{C^\alpha} := \sup_{t\in[0,1]} \|x_t\| + \|x\|_\alpha$$

is a norm, turning $C^\alpha([0,1],E)$ into a Banach space.

8.3.3 Assume that a path $x\colon [0,1] \to E$ with values in a Banach space is α-Hölder in the sense of Definition 8.12. If $\alpha > 1$, show that x is differentiable and constant.

8.3.4 Consider for $N \in \mathbb{N}$ the subset $\mathcal{T}_1^N(\mathbb{R}^d) = \{x \in \mathcal{T}^N(\mathbb{R}^d) \mid \pi_0(x) = 1\}$ (i.e. all elements in the tensor algebra whose zeroth degree term is 1) (recall that these are invertible!) set $\rho_N(x,y) := |x^{-1} \otimes y|$ where, as in (8.12), we have

$$|(1,x^1,x^2,\dots,x^n)| = \max_{k=1,\dots,N}\{(k!\|x^k\|)^{1/k}\} + \max_{k=1,\dots,N}\{(k!\|((x^{-1})^k\|)^{1/k}\}.$$

(a) Show that $|\cdot|$ is subadditive, that is, $|x \otimes x'| \le |x| + |x'|$.

(b) Given $\lambda \in \mathbb{R}$, define the dilation $\delta_\lambda\colon \mathcal{T}_1^N(\mathbb{R}^d) \to \mathcal{T}_1^N(\mathbb{R}^d)$, $(x^k)_{0\le k\le N} \mapsto (\lambda^k x^k)_{0\le k\le N}$. Show that $|\delta_\lambda(x)| = |\lambda||x|$.

(c) Show that ρ_N is a left-invariant metric on $G^N(\mathbb{R}^d)$, that is, $\rho_N(b\otimes x, b\otimes y) = \rho_N(x,y)$.

(d) Let $N \ge \lfloor 1/\alpha \rfloor$. Prove that a multiplicative functional $X\colon \Delta \to G^N(\mathbb{R}^d)$ is an α-rough path if and only if (8.13) is satisfied.
Hint: You will need Lemma 8.18 and can use (without a proof) that the kth component of $\mathbf{X}^{-1}$ arises by applying a linear map to the kth component of $\mathbf{X}$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

8.4 Rough Paths and the Shuffle Algebra

In this section we will broaden the scope of the investigation. The reason for this is twofold. On one hand, it allows us to give some further motivation to the question why infinite-dimensional differential geometry is of relevance in rough path theory. On the other hand, there is a pleasing framework which generalises the construction of the groups $G^\infty(\mathbb{R}^d)$ and leads to a whole class of infinite-dimensional Lie groups which have recently been found in a variety of mathematical contexts.

To start, consider again a smooth path $X\colon [0,1]\to\mathbb{R}^d$ with components $X^i, i=1,\dots d$. Now the iterated integrals comprising the signature $S_N(X)$ satisfy several algebraic conditions. For example, for $S_2(X)_{s,t} = (1, X_{s,t}, \int_s^t X_{s,r}\mathrm{d}X)$ the product rule of ordinary calculus yields

$$X^i_{s,t}\cdot X^j_{s,t} = \int_s^t X^i_{s,r}\mathrm{d}X^j + \int_s^t X^j_{s,r}\mathrm{d}X^i = \int_s^t\int_s^r \mathrm{d}X^i\mathrm{d}X^j + \int_s^t\int_s^r \mathrm{d}X^j\mathrm{d}X^i. \tag{8.14}$$

There are similar identities for higher-order iterated integrals and this is ultimately responsible for the non-linear set $G^N(\mathbb{R}^d)$ being the correct state space of a rough path. The question is of course how these identities and the associated combinatorics can be conveniently expressed. These questions lead to the so-called shuffle product.

8.21 Definition (Shuffle algebra (Reutenauer, 1993)) Consider the set $\mathcal{A} = \{1,2,\dots,d\}$, which in this context is called an *alphabet* and its elements *letters*. By concatenation we can construct words from the letters and the set $\mathcal{A}^*$ of all words including the empty word $\emptyset$. We construct now the shuffle algebra $\mathrm{Sh}(\mathcal{A}) = \mathbb{R}\mathcal{A}^*$ as the vector space generated by the words over $\mathcal{A}$. Note that, as a locally convex space, $\mathrm{Sh}(\mathcal{A})$ is isomorphic to the direct sum of countably many copies of $\mathbb{R}$; see Example A.36. Moreover, $\mathcal{A}^*$ and thus also $\mathrm{Sh}(\mathcal{A})$ is graded by word length, that is, if $w = a_1\cdots a_n \in \mathcal{A}^*$ for letters a_i, we set $|w| = n$ and say that an element of $\mathrm{Sh}(\mathcal{A})$ is homogeneous of degree $n\in\mathbb{N}_0$ if it is a linear combination of words of length n. For the algebra structure let $a,b\in\mathcal{A}$ and $u,w\in\mathcal{A}^*$. Then the *shuffle product* is defined recursively by

$$\emptyset \sqcup w = w \sqcup \emptyset = w, \quad (au)\sqcup(bw) := a(u\sqcup(bw)) + b((au)\sqcup w).$$

We will see in Exercise 8.4.1 that it extends to a continuous product on $\mathrm{Sh}(\mathcal{A})$, whence $\mathrm{Sh}(\mathcal{A})$ becomes a unital locally convex algebra, called the *shuffle algebra*.[7]

[7] The shuffle algebra is actually a graded bialgebra. The coproduct is given by deconcatenation of words

$$\delta\colon \mathrm{Sh}(\mathcal{A}) \to \mathrm{Sh}(\mathcal{A})\otimes\mathrm{Sh}(\mathcal{A}),$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

It turns out that the shuffle product encodes the combinatorics keeping track of the non-linear identities in the signature. For this we define for a smooth path $X\colon [0,1] \to \mathbb{R}^d$ with (truncated) signature $S_N(X), N \in \mathbb{N} \cup \{\infty\}$, the evaluation

$$\langle S_N(X), w\rangle := \int_s^t \int_s^{t_{n-1}} \cdots \int_s^{t_1} \mathrm{d}X^{a_1}\mathrm{d}X^{a_2}\cdots\mathrm{d}X^{a_m}$$

for $w = a_1\cdots a_n \in \mathcal{A}^*, |w| \le N$. With this notation in place, one can show (see Exercise 8.4.2) that the chain rule of ordinary calculus implies

$$\langle S_N(X)_{s,t}, w\rangle\langle S_N(X)_{s,t}, u\rangle = \langle S_N(X)_{s,t}, w \sqcup\!\sqcup u\rangle, \quad \text{for all } w,u \in \mathcal{A}^*. \tag{8.15}$$

Thus in view of (8.15), we see that the chain rule induces some non-linear constraints to the signature. Indeed, these constraints are responsible for the signature to take its values in $G^N(\mathbb{R}^d)$ instead of the full tensor algebra. The role these identities play will become clearer once we change our point of view slightly. For this we identify the kth level of the signature $S_\infty(X)$ as a mapping into $(\mathbb{R}^d)^{\otimes k} \cong \mathbb{R}^{d^k}$ whose components are given by the maps $\Delta \to \langle S_\infty(X), w\rangle$ for all $w \in \mathcal{A}^*$ with $|w| = k$. Hence for every pair $(s,t) \in \Delta$, we can identify the signature with a functional

$$\langle S_\infty(X)_{s,t}, \cdot\rangle\colon \mathrm{Sh}(\mathcal{A}) \to \mathbb{R}, \quad \mathcal{A}^* \ni w \mapsto \langle S_\infty(X)_{s,t}, w\rangle.$$

In Exercise 8.4.2 we will show that $\langle S_\infty(X)_{s,t}, \cdot\rangle$ is continuous, that is, it takes its values in the continuous dual $(\mathrm{Sh}(\mathcal{A}))' \cong \prod_{w\in\mathcal{A}^*}\mathbb{R}$ (this follows from Exercise B.1.2 as the dual of a locally convex direct sum of copies of $\mathbb{R}$ is the direct product; Meise and Vogt, 1997, Proposition 24.3). From this identification of the dual, one deduces at once that the signature gives rise to a continuous map $\langle S_\infty(X), \cdot\rangle\colon \Delta \to (\mathrm{Sh}(\mathcal{A}))'$. Elements in the image of $\langle S_\infty(X)_{s,t}, \cdot\rangle$ map the algebra product to the multiplication in $\mathbb{R}$ by (8.15), whence they are algebra morphisms.

Algebra morphisms from the shuffle algebra are called *characters* of the algebra $\mathrm{Sh}(\mathcal{A})$. We have already mentioned that the shuffle algebra carries more structure and is indeed a Hopf algebra. A *Hopf algebra* A is an algebra which

$$\delta(a_1a_2\cdots a_n) = \sum_{i=0}^{n} a_1a_2\cdots a_i \otimes a_{i+1}\cdots a_{i+2}\cdots a_n, \quad a_1,\dots,a_n \in \mathcal{A}.$$

As the subspace of homogeneous elements of degree 0 is 1-dimensional, the graded bialgebra $\mathrm{Sh}(\mathcal{A})$ becomes a Hopf algebra (see Manchon, 2008 for an introduction). This means that $\mathrm{Sh}(\mathcal{A})$ admits an antipode, that is, a mapping $S\colon \mathrm{Sh}(\mathcal{A}) \to \mathrm{Sh}(\mathcal{A})$ connecting the algebra and coalgebra structures and given by

$$S(a_1a_2\dots a_n) = (-1)^n a_n a_{n-1}\dots a_1 \quad \text{for all } a_1,\dots a_n \in \mathcal{A}.$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

is simultaneously a coalgebra $A \to A \otimes A$ such that the algebra and coalgebra structures are connected via the so-called antipode $S\colon A \to A$. We refer to Manchon (2008), for example, for the full definition and many examples. Now, for a Hopf algebra such as the shuffle algebra, the characters form a group $\mathcal{G}(\mathrm{Sh}(\mathcal{A}),\mathbb{R})$ under the convolution product

$$\varphi \star \psi(w) := \varphi \otimes \psi(\delta(w)), \text{ where } \delta \text{ is the deconcatenation coproduct.}$$

Note that we exploit here $\mathbb{R} \otimes \mathbb{R} \cong \mathbb{R}$. Inversion in the character group is given by precomposition with the antipode, that is, $\varphi^{-1} = \varphi \circ S$ (see e.g. Bogfjellmo et al., 2016, Lemma 2.3). Summing up, we have just proved the following.

8.22 Lemma *For a smooth path $X\colon [0,1] \to \mathbb{R}^d$, the signature induces a continuous map*

$$\langle S_\infty(X),\cdot\rangle\colon \Delta \to \mathcal{G}(\mathrm{Sh}(\mathcal{A}),\mathbb{R}) \subseteq \mathrm{Sh}(\mathcal{A})', \quad (s,t) \mapsto \langle S_\infty(X)_{s,t},\cdot\rangle$$

with values in the character group of the Hopf algebra $\mathrm{Sh}(\mathcal{A})$.

More generally, if we consider the kth level of an element

$$\mathbf{X} = (1, X^1, X^2, \ldots) \in G^\infty(\mathbb{R}^d)$$

we see that $X^k \in (\mathbb{R}^d)^{\otimes k}$ can be written as

$$X^k = \sum_{w \in \mathcal{A}^*, w = a_1 \cdots a_k} \alpha_{\mathbf{X}}(w) e_{a_1} \otimes \cdots \otimes e_{a_k},$$

where e_{a_ℓ} is the standard basis vector labelled by $a_\ell \in \mathcal{A} = \{1,\ldots,d\}$ and $\alpha_{\mathbf{X}}(w)$ some real coefficient. With some work on the algebra, one can prove the following.

8.23 Lemma (Ree, 1958, Theorem 2.6) *Let $\mathcal{A} = \{1,\ldots,d\}$ and for $\mathbf{X} = (1, X^1, X^2, \ldots)$ we write $X^k = \sum_{w \in \mathcal{A}^*, w = a_1 \cdots a_k} \alpha_{\mathbf{X}}(w) e_{a_1} \otimes \cdots \otimes e_{a_k}$. Then the following map is a well-defined group isomorphism:*

$$\Psi\colon G^\infty(\mathbb{R}^d) \to \mathcal{G}(\mathrm{Sh}(\mathcal{A}),\mathbb{R}), \quad \mathbf{X} = (1, X^1, X^2, \ldots) \mapsto \psi_{\mathbf{X}},$$
$$\text{where } \psi_{\mathbf{X}} \text{ is defined on } \mathcal{A}^* \text{ as } \psi_{\mathbf{X}}(w) = \alpha_{\mathbf{X}}(w) \quad \text{for all } w \in \mathcal{A}^*. \tag{8.16}$$

Moreover, a similar statement holds for the character group of the truncated shuffle algebra and the groups $G^N(\mathbb{R}^d), N \in \mathbb{N}$.

Now turning to the topological side, $G^\infty(\mathbb{R}^d)$ is a Lie group with respect to the subspace topology induced by the tensor algebra. As a locally convex space, the tensor algebra is isomorphic to a countable product of copies of the reals. We can also topologise the character group $\mathcal{G}(\mathrm{Sh}(\mathcal{A}),\mathbb{R}) \subseteq \mathrm{Sh}(\mathcal{A})' = \prod_{w \in \mathcal{A}^*} \mathbb{R}$ with the subspace topology. This topology even turns the character

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

group into an infinite-dimensional Lie group (Bogfjellmo et al., 2016, Theorem A). Thus we have the following additional statement.

8.24 Proposition *The group isomorphism* (8.16) *is an isomorphism of Lie groups. Thus there is a bijection between* α*-rough path* $\mathbf{X}$ *with values in* $G^\infty(\mathbb{R}^d)$ *and continuous maps*

$$\Psi_{\mathbf{X}} \colon \Delta \to \mathcal{G}(\mathrm{Sh}(\mathcal{A}), \mathbb{R}), \quad (t,s) \mapsto \psi_{\mathbf{X}_{s,t}}$$

such that for every $w \in \mathcal{A}^*$*, the component path* $\Delta \to \mathbb{R}$*,* $(t,s) \mapsto \psi_{\mathbf{X}_{s,t}}(w)$ *satisfies the '*$k\alpha$*-Hölder condition'*

$$\sup_{\substack{(s,t)\in\Delta \\ s\neq t}} \frac{|\psi_{\mathbf{X}_{s,t}}(w)|}{|t-s|^{k\alpha}} < \infty. \tag{8.17}$$

Proof In Exercise 8.4.3 you will show that (8.16) is the restriction of an isomorphism of locally convex spaces, hence an isomorphism of topological groups. Now by Proposition 8.7 the group $G^\infty(\mathbb{R}^d)$ is a Lie group and a submanifold of the tensor algebra. From Bogfjellmo et al. (2016, Theorem B) we know that also $\mathcal{G}(\mathrm{Sh}(\mathcal{A}), \mathbb{R})$ is a Lie group and a submanifold of the dual space of the shuffle algebra. Hence we deduce from Lemma 1.39 that (8.16) is already smooth and thus an isomorphism of Lie groups.

Every α-rough path $\mathbf{X}$ with values in $G^\infty(\mathbb{R}^d)$ is a continuous $G^\infty(\mathbb{R}^d)$-valued map (since its components are continuous in the product topology $G^\infty(\mathbb{R}^d) \subseteq \mathcal{T}^\infty(\mathbb{R}^d)$) and the kth level satisfies a $k\alpha$-Hölder condition. Hence $\Psi_{\mathbf{X}} := \Psi \circ \mathbf{X} \colon \Delta \to \mathcal{G}(\mathrm{Sh}(\mathcal{A}), \mathbb{R})$ is continuous. Up to changing the Hölder constant, we can change the norm on $(\mathbb{R}^d)^{\otimes k} \cong \mathbb{R}^{d^k}$ to the maximum norm. This shows that the components $\Psi_{\mathbf{X}}(w)$ inherit any Hölder condition the kth level of $\mathbf{X}$ satisfies, that is, (8.17) holds. Vice versa, if (8.17) holds for every $w \in \mathcal{A}^*$ with $|w| = k$, then the kth level of $\mathbf{X}$ satisfies a $k\alpha$-Hölder condition. This establishes the claimed bijection. □

Proposition 8.24 yields a nice interpretation of a (weakly geometric) rough path as a 'multiplicative functional' on the shuffle algebra. Of course this was not intended when T. Lyons originally coined the term 'multiplicative functional'.

8.25 Remark We are finally able to give a complete and easy proof of Lemma 8.18: In view of Proposition 8.24, a rough path $\mathbf{X}$ takes its values in $\mathcal{G}(\mathrm{Sh}(\mathcal{A}), \mathbb{R})$ and the group structure is isomorphic to the one of $G^\infty(\mathbb{R}^d)$. Hence the pointwise inverse $\mathbf{X}^{-1}$ can be computed via the group structure. Now it is a well-known fact about characters of a Hopf algebra that for a character ϕ, its inverse in the character group is given by $\phi \circ S$, where S is the antipode

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

of the Hopf algebra (Bogfjellmo et al., 2016, Lemma 2.3). Thus in the case of the shuffle Hopf algebra, this yields for an arbitrary linear combination of the words of length k the formula

$$S\left(\sum_{i_1,\dots i_k\in\{1,\dots,d\}} a_{(i_1,\dots,i_k)} i_1\cdots i_k\right) = \sum_{i_1,\dots i_k\in\{1,\dots,d\}} (-1)^k a_{(i_1,\dots,i_k)} i_k\cdots i_1.$$

Now the kth component of $\mathbf{X}$ corresponds to a linear combination of basis element $e_{i_1}\otimes e_{i_2}\otimes\cdots\otimes e_{i_k}$. The isomorphism $\mathcal{G}(\mathrm{Sh}(A),\mathbb{R})\cong G^\infty(\mathbb{R}^d)$ identifies $i_1 i_2\dots i_k$ with $e_{i_1}\otimes\cdots\otimes e_{i_k}$, whence the inversion formula shows the kth component of the pointwise inverse $\mathbf{X}^{-1}$ arises by permuting coefficients (and multiplying by $(-1)^k$). Hence the kth component of $\mathbf{X}^{-1}$ is $k\alpha$-Hölder if the kth component of $\mathbf{X}$ is so.

Exercises

8.4.1 Let $\mathcal{A}=\{1,\dots,d\}$ be a finite alphabet and $\mathcal{A}^*$ the monoid of words generated by $\mathcal{A}$. Recall that $\mathrm{Sh}(\mathcal{A})=\mathbb{R}\mathcal{A}^*$ is the vector space generated by $\mathcal{A}^*$. Show that:

(a) for all words $(u\sqcup\!\sqcup v)\sqcup\!\sqcup w = u\sqcup\!\sqcup(v\sqcup\!\sqcup w)$, $u\sqcup\!\sqcup v = v\sqcup\!\sqcup u$ and $|u\sqcup\!\sqcup v| = |u|+|v|$ hold;

(b) one can bilinearly extend the shuffle product to $\sqcup\!\sqcup$: $\mathrm{Sh}(\mathcal{A})\times\mathrm{Sh}(\mathcal{A})\to\mathrm{Sh}(\mathcal{A})$;

(c) the box-topology from Example A.36 makes $\sqcup\!\sqcup$: $\mathrm{Sh}(\mathcal{A})\times\mathrm{Sh}(\mathcal{A})\to\mathrm{Sh}(\mathcal{A})$ continuous and $(\mathrm{Sh}(\mathcal{A}),\sqcup\!\sqcup)$ a locally convex algebra.
Hint: By Exercise B.1.2 it suffices to show that for every pair of words v,w the shuffle induces a continuous map $\mathbb{R}\times\mathbb{R}\to\mathrm{Sh}(\mathcal{A})$, $(a,b)\mapsto(av)\sqcup\!\sqcup(bw)$.

8.4.2 Let $X\colon[0,1]\to\mathbb{R}^d$ be a smooth path with signature $S_N(X)$ (for some $N\in\mathbb{N}\cup\{\infty\}$). Let $w=a_1\cdots a_n\in\mathcal{A}^*$. Define $\langle S_N(X)_{s,t},w\rangle = \int_s^t \mathrm{d}X^{a_1}\mathrm{d}X^{a_2}\cdots\mathrm{d}X^{a_n}$. Show that:

(a) (8.15) holds, that is, $\langle S_N(X)_{s,t},w\rangle\langle S_N(X)_{s,t},u\rangle=\langle S_N(X)_{s,t},w\sqcup\!\sqcup u\rangle$, for all $w,u\in\mathcal{A}^*$. *Hint:* Use induction by word length and the chain rule. For $|w|=1$, the claim becomes (8.14).

(b) For every pair $(s,t)\in\Delta$, the linear map $\langle S_\infty(X)_{s,t},\cdot\rangle\colon\mathrm{Sh}(\mathcal{A})\to\mathbb{R}$ is continuous, where we endow the shuffle algebra with the box-topology from Example A.36.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(c) The signature gives rise to a continuous map $\langle S_\infty(X), \cdot \rangle \colon \Delta \to (\mathrm{Sh}(\mathcal{A}))' \cong \prod_{w \in \mathcal{A}^*} \mathbb{R}$. *Hint:* The identification is given by $\mathrm{Sh}(\mathcal{A}) \ni f \mapsto (f(w))_{w \in \mathcal{A}^*} \in \prod_{w \in \mathcal{A}^*} \mathbb{R}$.

8.4.3 Show that there is a canonical isomorphism $\mathrm{Sh}(\mathcal{A})' = \prod_{n \in \mathbb{N}} \mathbb{R} \cong \mathcal{T}^\infty(\mathbb{R}^d)$ of locally convex spaces if $\mathcal{A} = \{1, \dots, d\}$. Deduce that this restricts to (8.16), whence (8.16) is a morphism of topological groups if we endow $G^\infty(\mathbb{R}^d)$ with its Lie group topology and $\mathcal{G}(\mathrm{Sh}(\mathcal{A}), \mathbb{R})$ with the subspace topology induced by embedding it in $\mathrm{Sh}(\mathcal{A})'$.

8.5 The Grand Geometric Picture (Rough Paths and Beyond)

In the last section we have seen that rough paths can be understood as certain continuous paths with values in the character group of the shuffle Hopf algebra. We will now extend the focus slightly to different types of rough paths. For this, recall that the signature of a smooth path had to satisfy certain algebraic identities which connected it to the shuffle algebra and its character group. For smooth paths this was a consequence of the chain rule of ordinary calculus. However, there are also many interesting objects, which do not satisfy a classical chain rule. One example for this is the Itô-integral from stochastic calculus (nevertheless we constructed the rough path Itô-lift of Brownian motion in Example 8.17). However, this motivates the idea of relaxing the requirement of (8.15) and recording iterated integrals of a path $X \colon [0,1] \to \mathbb{R}^d$ of the form

$$\int_s^t \left(\int_s^r \mathrm{d}X^i \right) \left(\int_s^r \mathrm{d}X^j \right) \mathrm{d}X^k. \tag{8.18}$$

For the rough paths we have considered so long, the iterated integral (8.18) (respectively the corresponding combination of levels of the rough path) simplifies via the shuffle identity to an iterated integral (resp. level of the rough path) we have already recorded. Relaxing the requirement (8.15) necessitates that we record objects corresponding to (8.18) as additional information. Following this approach, one arrives at Gubinelli's concept of a *branched rough path* (Gubinelli, 2010). Branched rough paths satisfy different algebraic identities than the rough paths we have been considering (which are called *weakly geometric rough paths* in the literature (Friz and Hairer, 2020)). By dropping (8.15), branched rough paths will not take values in the character group of the shuffle algebra. However, this can easily be remedied by replacing the shuffle algebra with another Hopf algebra. This leads us to the following concept.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

8.26 Definition Let $\mathcal{H}$ be a graded and connected Hopf algebra[8] and let $(\mathcal{G}(\mathcal{H},\mathbb{R}),\star)$ be its character group. Fix a basis B of the vector space $\mathcal{H}$ consisting of homogeneous elements. Then an α-rough path over $\mathcal{H}$ is a map

$$\mathbf{X}\colon \Delta \to \mathcal{G}(\mathcal{H},\mathbb{R})$$

which satisfies Chen's relation

$$\mathbf{X}_{s,u} \star \mathbf{X}_{u,t} = \mathbf{X}_{s,t}, \qquad s \le u \le t \in [0,1],$$

and for every $h \in B$ we have the graded Hölder condition $|\langle \mathbf{X}_{s,t}, h\rangle| \lesssim |t-s|^{\alpha|h|}$ (where $\langle\cdot,h\rangle$ is evaluation in h). If $\dim\mathcal{H}_1 = d$ and $B \cap \mathcal{H}_1 = \{e_1,\dots,e_d\}$, an $\mathcal{H}$-rough path $\mathbf{X}$ *lifts* $x\colon [0,1] \to \mathbb{R}^d, x(t) = (x_1(t),\dots,x_d(t))$ if $\langle \mathbf{X}_{s,t}, e_i\rangle = x_i(t) - x_i(s), i = 1,\dots,d$ for all $s,t \in \Delta$.

Since the Hopf algebra $\mathcal{H}$ is graded, one can also define truncated versions of the rough paths by looking at the character group of truncated Hopf algebra. Note that the choice of basis B is part of the definition of an $\mathcal{H}$-rough path and the graded Hölder condition will depend on the choice of grading and the base.

8.27 Example The concept of an α-rough path over $\mathcal{H}$ is quite flexible and we illustrate this with the following list of different types of rough paths appearing in the literature:

- For $\mathcal{H} = \mathrm{Sh}(\mathcal{A})$ and $B = \mathcal{A}^*$, we recover the notion of a (weakly) geometric rough path as discussed in §§8.3 and 8.4.
- If $\mathcal{H}$ is the Butcher–Connes–Kreimer Hopf algebra of (decorated) rooted trees (Bogfjellmo et al., 2016, Example 4.6) and B the base of all forests, one obtains branched rough paths (Gubinelli, 2010).
- For the Munthe-Kaas–Wright Hopf algebra and B the base consisting of planar forests, one obtains the notion of a planarly branched rough path. These objects generalise the concept of a branched rough path to homogeneous spaces. See Curry et al. (2020) for more information.

While switching the Hopf algebra allows us to treat different concepts of rough paths, the general geometric theory stays the same for the different types of rough paths which can be treated in this generalised setting. For example, there is a generalised version of the Lyons–Victoire lifting theorem (see Tapia and Zambotti, 2020, Theorem 3.4) and the character groups can be endowed with an infinite-dimensional Lie group structure for which every rough path becomes a continuous map (see Bogfjellmo et al., 2016, Theorem A). Thus the

[8] Graded and connected means that $\mathcal{H} = \bigoplus_{n\in\mathbb{N}_0} \mathcal{H}_n$ as vector spaces with $\mathcal{H}_0 \cong \mathbb{R}$ and the algebra and coalgebra structures are compatible with the grading; see Bogfjellmo et al. (2016). Recall that an element $x \in \mathcal{H}$ is called homogeneous of degree $|x| = n$ if $x \in \mathcal{H}_n$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

theory for weakly geometric rough paths from §8.4 carries over to the generalised concept of rough paths over $\mathcal{H}$. However, there are also new geometric features which appear in this generalised setting. For example, the following result hints at interesting geometry:

8.28 Proposition (Tapia and Zambotti, 2020, Theorem 1.2) *For every $\alpha \in]0,1[$ with $\alpha^{-1} \notin \mathbb{N}$ the α-branched rough paths form a homogeneous space under the action of a vector space of α-Hölder functions.*

The term 'homogeneous space' in Proposition 8.28 means that there is a transitive and free group action $G \times X \to X$ of a vector space G of Hölder continuous functions on the space of branched rough paths X. The canonical manifold structures on the spaces and groups should turn the branched rough paths into an infinite-dimensional homogeneous space in the sense of differential geometry (see Example 3.55); however, this question was, to the best of my knowledge, not yet investigated. While there is a lot more which could be said about rough paths and their interplay with (finite- and infinite-dimensional) geometry we will not go into more detail here. Instead, let us briefly point out a general theme underlying the idea to identify rough paths as paths into the character group of a suitable Hopf algebra.

8.29 (Elements in the character group as formal power series) Assume that we have a graded Hopf algebra $\mathcal{H} = \bigoplus_{n\in\mathbb{N}_0} \mathcal{H}_n$ with B a vector space base of $\mathcal{H}$ consisting only of homogeneous elements. Instead of thinking of an element in the (continuous) dual space $\mathcal{H}' = \prod_{n\in\mathbb{N}_0} \mathcal{H}_n'$ as a linear map, we can identify it with its values on the base B and write it as a *formal power series*

$$\psi = \sum_{w\in B} \langle \psi, w\rangle w, \qquad \text{with } \langle \psi, w\rangle := \psi(w). \tag{8.19}$$

Now elements in the character group $\mathcal{G}(\mathcal{H},\mathbb{R})$ are algebra morphisms. Hence it suffices to record their values on a set C which generates the algebra $\mathcal{H}$ in the sense that $\mathcal{H}$ is isomorphic (as an algebra) to the (not necessarily commutative) polynomial algebra $\mathbb{R}\langle C\rangle$. If we have chosen a set C generating $\mathcal{H}$ in this sense, it suffices to record, instead of the formal power series (8.19), the series

$$\psi = \sum_{w\in C} \langle \psi, w\rangle w, \qquad \text{if } \psi \in \mathcal{G}(\mathcal{H},\mathbb{R}). \tag{8.20}$$

8.30 Example For the shuffle algebra $\mathrm{Sh}(\mathcal{A})$, we can represent a functional as the formal power series

$$\psi = \sum_{w\in\mathcal{A}^*} \langle \psi, w\rangle w.$$

These series are also called *word series* and they are studied in relation to dynamical systems and numerical integration (Murua and Sanz-Serna, 2017).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Note that the shuffle algebra can be interpreted as a commutative polynomial algebra generated by the Lyndon words (Reutenauer, 1993, Chapter 5).

While 8.29 seems to be only a notational trick to write a functional as a series, this idea leads to a rich source of examples. The idea one should have when looking at the series expansions (8.19) and (8.20) is that these series represent some kind of Taylor expansion.[9] It is then the role of the Hopf algebra $\mathcal{H}$ to describe the combinatorics for the objects which define the Taylor expansion. Groups of (formal) power series modelled as characters on suitable Hopf algebras arise in a variety of contexts. Beyond rough paths here one should mention the following applications: Hairer's regularity structures for stochastic partial differential equations (Friz and Hairer, 2020, Chapter 13), word series in numerical analysis (Murua and Sanz-Serna, 2017), Chen–Fliess series in control theory (Gray et al., 2022) and the Connes–Kreimer approach to the renormalisation of quantum field theories (Manchon, 2008). In these applications, the differential geometry of the character group is of interest. As a concrete example, recall the connection to numerical integrators.

8.31 Example Consider the (time-independent) ordinary differential equation

$$y'(t) = F(y), \text{ where } F\colon \mathbb{R}^d \to \mathbb{R}^d \text{ is a vector field.} \tag{8.21}$$

Assume that we are trying to compute a power series solution to (8.21). Then we compute derivatives of y via the chain rule:

$$y'=F(y), y''=dF(y;F(y)), y'''=d^2F(y;F(y),F(y))+dF(y;dF(y;F(y))),\ldots.$$

Fixing a starting point y_0 we only need iterated differentials of F and information on where derivatives were inserted in the arguments of the differential. The combinatories can be handled by encoding the information via rooted trees, that is, as finite graphs with a special node called the root (displayed as the nethermost node in the following):

Formally, we write $\emptyset$ for the empty tree. Every rooted tree τ can be written recursively as $\tau = [\tau_1,\ldots,\tau_m]$, where the τ_i are trees whose roots are grafted to a common new root. For example, $[\bullet,\bullet]$ = . Then the iterated differentials of F can recursively be encoded via the *elementary differentials* defined as $E_F(\bullet, y) := F(y)$, and for a rooted tree τ we set

[9] For rough paths this is most easily seen in the context of controlled rough paths; see Friz and Hairer (2020, Section 4.6).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$E_F(\tau)(y) \coloneqq d^n F(y; E_F(\tau_1, y), \dots, E_F(\tau_n, y)) \text{ for } \tau = [\tau_1, \dots, \tau_n].$$

This leads to a formal power series, called a *B-series*

$$B_F(\psi, y, h) \coloneqq y + \sum_{\tau \text{ rooted tree}} \frac{h^{|\tau|}}{\sigma(\tau)} \psi(\tau) E_F(\tau, y), \tag{8.22}$$

where $h \in \mathbb{R}, a(\tau) \in \mathbb{R}$, $|\tau|$ is the number of nodes in the tree τ and $\sigma(\tau)$ is a certain symmetry factor. B-series model certain numerical solutions to (8.21), where the parameter h is interpreted as the step size of the numerical method. One identifies the B-series (8.22) with its coefficients $\{\psi(\tau)\}_{\tau \text{ rooted tree}}$. As the rooted trees generate the Butcher–Connes–Kreimer Hopf algebra, ψ extends to a character of this Hopf algebra. The character group of the Butcher–Connes–Kreimer Hopf algebra is in numerical analysis known as the *Butcher group* (Bogfjellmo and Schmeding, 2017). Its elements encode numerical integration schemes such as Runge–Kutta methods and the group product models composition of schemes.

From the perspective of numerical analysis, the Butcher group is a convenient tool as its algebraic and differential geometric structure is of interest in the analysis of numerical integrators. Unfortunately, it is also a very large group with many elements which do not correspond to any (locally) convergent integration scheme. This leads to subgroups which admit a stronger topology turning them into Lie groups while still containing all elements of interest (Dahmen and Schmeding, 2020). In the context of rough paths, a similar problem is the question of whether there is a differentiable structure on the subgroup generated by all α-rough paths for some fixed $\alpha \in]0, 1[$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Appendix A
A Primer on Topological Vector Spaces and Locally Convex Spaces

This section contains some auxiliary results on topological vector spaces and locally convex spaces in particular. Most of the results are standard and can be found in textbooks such as Meise and Vogt (1997), as well as Jarchow (1981). Note that for some of the results (e.g. Proposition A.4) in this appendix, it is essential that we only consider Hausdorff topological vector spaces. Since we are only working with real vector spaces, some of the proofs simplify substantially (compare to the general proofs for $\mathbb{R}$ and $\mathbb{C}$; see Rudin (1991, Chapter 1)).

A.1 Basic Material on Topological Vector Spaces

A vector space with a Hausdorff topology making vector addition and scalar multiplication continuous is called a topological vector space or TVS (see Definition 1.1). Note that a morphism of TVS is a continuous linear map. In particular, two TVS are isomorphic (as TVS) if they are isomorphic as vector spaces and the isomorphism is a homeomorphism.

Conventions Let U, V be subsets of a (topological) vector space E, $s \in \mathbb{R}$ and $I \subseteq \mathbb{R}$. Then we define

$$U + V := \{z = u + v \mid u \in U, v \in V\}, \quad sU := \{z = su \mid u \in U\}, \quad I \cdot U := \bigcup_{s \in I} sU.$$

A.1 Definition Let $(E, \mathcal{T})$ be a TVS and U a subset of E. We say that U is

(a) a 0-*neighbourhood* if U is a neighbourhood of 0;
(b) *bounded* if for every 0-neighbourhood V there is $s > 0$ with $U \subseteq sV$.

In general, topological vector spaces sequences are not sufficient to test, for example, continuity. Instead one would need nets to test for continuity and a

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

complete topological vector space should be defined in terms of convergence of Cauchy nets (see Meise and Vogt, 1997, p. 258). However, for our calculus we usually do not need complete spaces and the limits we consider can always be described in terms of sequential limits. Thus we do not go into details here and stay in the realm of the more familiar sequences. A sequence $(x_n)_{n\in\mathbb{N}} \subseteq E$ is called

(c) a *Cauchy sequence* if for every 0-neighbourhood $V \subseteq E$ there exists $N \in \mathbb{N}$ such that
$$x_n - x_m \in V \quad \text{for all } n, m \geq N;$$

(d) a *Mackey–Cauchy sequence* if there exists a bounded subset $B \subseteq E$ and a family $m_{k,l} \in \mathbb{N}$ for $k, l \in \mathbb{N}$ such that
$$m_{k,l}(x_k - x_l) \in B, \quad \text{for all } k, l \in \mathbb{N}$$
and such that for every $R > 0$ there is $N \in \mathbb{N}$ with $m_{k,l} > R$ if $k, l > N$ (i.e. $m_{k,l} \to \infty$). Note that every Mackey–Cauchy sequence is a Cauchy sequence.

Now we say that the topological vector space $(E, \mathcal{T})$ is

(e) *sequentially complete* if every Cauchy sequence in E converges;
(f) *Mackey complete* if every Mackey–Cauchy sequence in E converges.

Mackey completeness as per (f) can be shown (see Kriegl and Michor, 1997, Theorem 2.14) to be equivalent to the notion from Definition 1.12.

A.2 Lemma *Let E be a topological vector space and $U \subseteq E$ a 0-neighbourhood. Then the following holds:*

(a) For each $x \in E$ the translation $\lambda_x : E \to E$, $y \mapsto x + y$ is a homeomorphism.
(b) For each $r \in \mathbb{R} \setminus \{0\}$, scaling $s_r : E \to E$, $x \mapsto rx$ is a homeomorphism.
(c) U contains a balanced *0-neighbourhood V, that is, $tV \subseteq V$ for each $|t| \leq 1$.*
(d) U contains a 0-neighbourhood W such that $W + W \subseteq U$.
(e) If $\mathcal{B}$ is a basis of 0-neighbourhoods, then for each $x \in E$ the set $\{x + W \mid W \in \mathcal{B}\}$ is a basis of x-neighbourhoods.
(f) Each 0-neighbourhood contains a closed 0-neighbourhood.
(g) If $K \subseteq E$ is compact and $U \subseteq E$ with $K \subseteq U$, then there exists $0 \in W \subseteq E$ such that $K + W \subseteq U$.

Proof **(a–b) and (e)**. The maps λ_x and s_r have inverses λ_{-x} and $s_{1/r}$. Thus the claim is clear from the definition of topological vector spaces. Since translations are homeomorphisms (a) implies (e).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(c). By continuity of the scalar multiplication $\mu\colon \mathbb{R}\times E\to E$, the set $\mu^{-1}(U)$ is open. Thus we can find $(-\varepsilon,\varepsilon)\times W\subseteq \mu^{-1}(U)$ and thus $V := (-\varepsilon,\varepsilon)W = \mu((-\varepsilon,\varepsilon)\times W)\subseteq U$. Then $[-1,1]V = V$ and V is a balanced 0-neighbourhood contained in U.

(d). As addition $\alpha\colon E\times E\to E$ is continuous with $\alpha(0,0)=0$, the preimage $\alpha^{-1}(U\times U)$ is a $(0,0)$-neighbourhood. We can thus find $W_1, W_2\subseteq E$ 0-neighbourhoods in E such that $W_1\times W_2\subseteq \alpha^{-1}(U)$. Then $W := W_1\cap W_2\subseteq U$ satisfies $W+W\subseteq U$.

(f). We conclude from (c) and (d) that there is a 0-neighbourhood V with $V-V\subseteq U$. For $w\in\overline{V}$ (closure), the set $w+V$ is a w-neighbourhood (by (e)). Hence we can pick $v_1\in V$ such that $v_1\in w+V$, that is, $v_1 = w+v_2$ for some $v_2\in V$. But then $w = v_1 - v_2\in U$, and so $\overline{V}\subseteq U$.

(g). For every $x\in K$ we can pick by (e) a 0-neighbourhood V_x such that $x+V_x\subseteq U$. By (d) there is $0\in W_x\subseteq E$ with $W_x+W_x\subseteq V_x$. Then $(x+W_x)_{x\in K}$ is an open cover of K and by compactness we can choose a finite subset $F\subseteq K$ with $K\subseteq\bigcup_{x\in F}(x+W_x)$. Then $W := \bigcap_{x\in F} W_x$ is an open 0-neighbourhood. For $y\in K$, there exists $x\in F$ such that $y\in x+W_x$. Then $y+W\subseteq x+W_x+w\subseteq x+V_x\subseteq U$. As y was arbitrary $K+W\subseteq U$. □

A.3 Proposition (Rudin, 1991, I Theorem 1.22) *If E is a topological vector space which contains a compact 0-neighbourhood, then E is finite dimensional.*

A.4 Proposition (Uniqueness of topology (Treves, 2006, Theorem 9.1)) *If E is a finite-dimensional topological vector space of dimension d, then $E\cong\mathbb{R}^d$ as topological vector spaces, where $\mathbb{R}^d$ carries the usual norm topology.*

A.5 Lemma *Let $f\colon E\to F$ be a linear map between topological vector spaces. Then f is continuous (open) if and only if it is continuous (open) in 0.*

Proof Clearly the conditions are necessary. To prove sufficiency, we assume that f is continuous in 0. Pick $x\in E$ and observe that the translations $\tau_{-x}(y) = y-x$ and $\tau_{f(x)}(z) = z+f(x)$ are continuous and even homeomorphisms. Thus $f(y) = \tau_{f(x)}\circ f\circ\tau_{-x}$ is continuous in x. Since x was arbitrary, f is continuous in every point.

Now let f be open in 0 and U an open set. For $x\in U$, $\tau_{-x}(U)$ is an open 0-neighbourhood, whence $\tau_{f(x)}\circ f\circ\tau_{-x}(U) = f(U)$ is open. □

En route towards locally convex spaces let us first recall some results on convex sets.

A.6 Definition A subset S of a topological vector space E is said to be

- *absorbent* if $E=\bigcup_{t>0} tS$;

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

- *convex* if for all $x, y \in S$ and $t \in [0,1]$, the linear combination $tx + (1-t)y \in S$;
- a *disc* if it is convex and balanced, that is, for all $x \in S$, $|\lambda| \leq 1, \lambda \in \mathbb{R}$, $\lambda x \in S$.

A.7 Lemma *Every 0-neighbourhood of a topological vector space is absorbent.*

Proof Let U be a 0-neighbourhood of the topological vector space E, and let $x_0 \in E$. Since scalar multiplication is continuous there exists some neighbourhood V of x_0 and $\delta > 0$ such that for all $x \in V$, $\lambda x \in U$ when $|\lambda| < \delta$, $\lambda \in \mathbb{R}$. Especially $\lambda x_0 \in U$, and thus $x_0 \in \lambda^{-1}U$ for $|\lambda^{-1}| > 1/\delta$. □

A.8 Example The ball $B_r(x) := \{y \in E \mid \|x - y\| < r\}$ in a normed space $(E, \|\cdot\|)$ is convex (and a disc if $x = 0$).

A.9 Lemma *The interior A° of a convex set A is convex.*

Proof Let $x, y \in A^\circ$. By Lemma A.2(c) there is some balanced 0-neighbourhood U such that $x + U \subseteq A^\circ$ and $y + U \subseteq A^\circ$ are neighbourhoods of x and y contained in A. For any $z = tx + (1-t)y$, $t \in [0,1]$ and $u \in U$ we have

$$z + u = tx + (1-t)y + tu + (1-t)u = t(x+u) + (1-t)(y+u).$$

As A is convex, $z + u \in A$ and thus $z + U \subseteq A$. Hence $tx + (1-t)y \in A^\circ$, for all $t \in [0,1]$. □

A.10 Lemma *If N is a convex 0-neighbourhood in a topological vector space, then N contains an open disc.*

Proof Consider first the set $M := -N \cap N$. If $|\lambda| \leq 1$ we see that $\lambda M = (-\lambda)N \cap \lambda N$. Now, as $0 \in N$ and N is convex, we have that $-\lambda N, \lambda N \subseteq N$. In particular, $\lambda M \subseteq M$ for all $|\lambda| \leq 1$, that is, M is balanced. By Lemma A.2(c), we can find a balanced 0-neighbourhood $U \subseteq N$. As U is balanced we see that $U \subseteq -N \cap N$. Hence the interior V of $M = -N \cap N$ is a convex 0-neighbourhood (by Lemma A.9 as it is the interior of an intersection of convex sets; Exercise A.1.3). Now the interior of a balanced set is again balanced (Exercise A.1.1), whence V is a disc. □

Exercises

A.1.1 Let B be a balanced subset of a topological vector space. Show that then also the interior of B is balanced.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

A.1.2 Let $U \subseteq E$ be a bounded 0-neighbourhood in a topological vector space E. Show that every 0-neighbourhood contains a set of the form $\{rU\}_{r\in]0,\infty[}$.
Hint: Use Lemma A.7 together with the fact that U is bounded.

A.1.3 Show that the intersection $C \cap D$ of two convex sets C and D is convex.

A.2 Seminorms and Convex Sets

In the main text we have defined locally convex spaces using seminorms. In this section we shall review seminorms and, in particular, their connection to convex sets (thus justifying the name 'locally convex space').

A.11 Definition A family $\mathcal{P}$ of seminorms on a vector space E is said to be *separating* if for each $x \in E$, $p(x) \neq 0$ for at least one $p \in \mathcal{P}$.

A.12 Proposition *Let E be a vector space and $(p_i)_{i\in I}$ a separating family of seminorms on E. Then a Hausdorff vector topology is generated by the subbase*

$$B_{i,\epsilon}(x_0) := \{x \in E \mid p_i(x - x_0) < \epsilon\}, \quad i \in I, \epsilon > 0, x_0 \in E. \tag{A.1}$$

Thus $(E, \{p_i\}_I)$ is a locally convex space and the topology contains a 0-neighbourhood basis of convex sets. Finally, each p_i is continuous with respect to the locally convex topology.

Proof Let us first note that the subbase (A.1) generates the initial topology induced by the family $\{q_i : E \to E/\ker p_i\}_{i\in I}$, where the right-hand side is endowed with the normed topology induced by p_i (see Exercise A.2.1). Let U be a 0-neighbourhood. Then U contains some finite intersection $\bigcap B_{i,\epsilon}(x)$, which is convex since the seminorm balls are convex and intersections preserve convexity; Exercise A.1.3. Thus every 0-neighbourhood contains a convex 0-neighbourhood. For the Hausdorff property we choose for $x_1, x_2 \in E$ a seminorm p_i such that $0 < p_i(x_1 - x_2)$. Set $\delta = p_i(x_1 - x_2)/3$. Now if $z \in B_{i,\delta}(x_2) \cap B_{i,\delta}(x_1)$ were non-empty, we must have

$$0 < \delta = p_i(x_1 - x_2) \leq p_i(x_1 - z) + p_i(z - x_2) \leq \frac{2}{3}\delta,$$

which is absurd. Therefore the intersection must be empty and E is Hausdorff.

We have continuity of addition since for each

$$U_{x+y} = B_{i_1,\epsilon_1}(x_0 + y_0) \cap B_{i_2,\epsilon_2}(x_0 + y_0) \cap \cdots \cap B_{i_n,\epsilon_n}(x_0 + y_0),$$

the neighbourhoods

$$U_{x_0} = B_{i_1,\epsilon_1/2}(x_0) \cap B_{i_2,\epsilon_2/2}(x_0) \cap \cdots \cap B_{i_n,\epsilon_n/2}(x_0)$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

and

$$U_{y_0} = B_{i_1,\epsilon_1/2}(y_0) \cap B_{i_2,\epsilon_2/2}(y_0) \cap \cdots \cap B_{i_n,\epsilon_n/2}(y_0)$$

satisfy $U_x + U_y \subset U_{x+y}$ since for any $x + y \in B_{i,\epsilon/2}(x) + B_{i,\epsilon/2}(y)$,

$$p_i(x + y) \leq p_i(x) + p_i(y) < \epsilon/2 + \epsilon/2 = \epsilon.$$

For the continuity of scalar multiplication, consider the neighbourhood

$$U_{\lambda_0 x_0} = B_{i_1,\epsilon_1}(\lambda_0 x_0) \cap B_{i_2,\epsilon_2}(\lambda_0 x_0) \cap \cdots \cap B_{i_n,\epsilon_n}(\lambda_0 x_0).$$

Let $|\lambda - \lambda_0| < \delta$ and $x \in B_{i,\delta}(x_0)$, then

$$\begin{aligned} p_i(\lambda x - \lambda_0 x_o) &= p_i((\lambda - \lambda_0)x + \lambda_0(x - x_0)) \\ &\leq |\lambda - \lambda_0| p_i(x) + |\lambda_0| p_i(x - x_0) \\ &< \delta(p_i(x - x_0) + p_i(x_0)) + |\lambda_0| p_i(x - x_0) \\ &< \delta(\delta + p_i(x_0) + |\lambda_0|) < \epsilon \end{aligned}$$

if δ is small enough. So we can find $\delta_1, \ldots, \delta_n$ and $\delta = \min\{\delta_1, \ldots, \delta_n\}$ such that

$$(\lambda, x) \in]\lambda_0 - \delta, \lambda_0 + \delta[\times U_{x_0} \implies \lambda x \in U_{\lambda_0 x_0}.$$

To see that p_i is continuous, let $]a, b[$ be an open interval in $[0, \infty[$, then $p_i^{-1}(]a, b[) = (E \setminus \overline{B}_{i,a}(0)) \cap B_{i,b}(0)$ is open being a finite intersection of open sets, when $\overline{B}_{i,a}(0) := \{x \in E \mid p_i(x) \leq a\}$. Since $p_i^{-1}([0, b[) = B_{i,b}(0)$, we deduce that p_i is continuous. □

A.13 Definition Let $(p_i)_{i \in I}$ be a family of seminorms. We say the family

- satisfies the *basis condition* if for each two seminorms p_i and p_j, there exists a third seminorm p_k and $C > 0$ such that

$$\max\{p_i(x), p_j(x)\} \leq C p_k(x), \quad \text{for all } x \in E;$$

- is called a *fundamental system* of seminorms, if it generates the topology on E and satisfies the basis condition.

Note that for a fundamental system of seminorms, the subbase (A.1) is a basis for the topology it generates.

A.14 Example Consider again the space of smooth functions $C^\infty([0, 1], \mathbb{R})$ with the Fréchet topology induced by the seminorms

$$\|f\|_n := \sup_{0 \leq k \leq n} \sup_{x \in [0,1]} \left| \frac{\mathrm{d}^k}{\mathrm{d}x^k} f(x) \right|.$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

For two of these seminorms we obviously have

$$\max\{\|f\|_n, \|f\|_m\} \leq \|f\|_{\max\{n,m\}} \text{ for all } f \in C^\infty([0,1],\mathbb{R}).$$

Hence these seminorms form a fundamental system of seminorms and their r-balls form a basis of the topology called the compact open C^∞-topology.

Let us associate now to every disc a seminorm. The upshot will be that one can equivalently define a locally convex space as a topological vector space with a 0-neighbourhood base consisting of convex sets.

A.15 Definition For a vector space E and a disc A in E, define the *Minkowski functional*, $p_A\colon E \to \mathbb{R}, p_A(x) := \inf\{t > 0. \mid x \in tA\}$, where $\inf \emptyset := \infty$.

A.16 Lemma *If U is a disc 0-neighbourhood in the locally convex space E, then the Minkowski functional p_U is a continuous seminorm on E.*

Proof By Lemma A.7, $p_U(x) \in [0,\infty[$ for all $x \in E$. For the triangle inequality, let $x, y \in E$, then if $x \in tU$ and $y \in sU$,

$$\frac{1}{t+s}(x+y) = \frac{t}{t+s}\frac{x}{t} + \frac{s}{t+s}\frac{y}{s} \in U$$

or rather $x + y \in (t+s)U$. Therefore, $p_U(x+y) \leq t + s = p_U(x) + p_U(y)$. Scalar factors can be taken out of the seminorm due to U being balanced: If $\lambda \in \mathbb{R} \setminus \{0\}$, $\lambda x = \frac{\lambda}{|\lambda|}|\lambda| x \in tU$ if and only if $|\lambda| x = \frac{|\lambda|}{\lambda}\lambda x \in tU$. Therefore, $p_U(\lambda x) = \inf\{t > 0 : \lambda x \in tU\} = \inf\{t > 0 : |\lambda| x \in tU\} = |\lambda| p_U(x)$. Continuity of the seminorm follows from $p_U^{-1}(]a,b[) = (E \setminus (\overline{aU})) \cap \bigcup_{0 \leq t \leq b} tU$ is open for all $a, b \in [0,\infty[$. □

As Lemma A.10 shows, every convex 0-neighbourhood gives rise to a disc and these give rise to seminorms by Lemma A.16. Thus an equivalent definition of a locally convex space (fitting the name better; see Rudin, 1991, Theorem 1.34 and Remark 1.38) is the following.

A.17 Definition A Hausdorff topological vector space is a locally convex space if it contains a 0-neighbourhood basis of convex sets.

Finally, let us recall Kolmogorov's normability criterion, which gives a (necessary and sufficient!) condition for a topological vector space to be normable, that is, the vector topology coincides with the topology induced by some norm.

A.18 Theorem (Kolmogorov's normability criterion) *A topological vector space E is normable if and only if E has a bounded and convex 0-neighbourhood.*

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Proof The criterion is necessary as in a normed space, the unit ball is a convex bounded 0-neighbourhood.

Now let E be a topological vector space with a bounded and convex 0-neighbourhood N. By Lemma A.10 we can pick an open disc $U \subseteq N$. Note that U is also bounded and define $\|x\| := p_U(x)$ for $x \in E$, where p_U is the Minkowski functional associated to U. Now, thanks to Exercise A.1.2, the sets sU, $s \in]0,\infty[$ form a neighbourhood base of 0 in E. If $x \neq 0$ is an element of E, the Hausdorff property implies that there exists $s > 0$ with $x \notin sU$, that is, $\|x\| \geq s > 0$. We deduce from Lemma A.16 that $\|\cdot\|$ is a (continuous) norm on E. Hence the norm topology induced by $\|\cdot\|$ is coarser than the original topology. Conversely, recall that since U is open, we have $\{x \in E \mid \|x\| < s\} = sU$. As the sU form a neighbourhood base, this shows that the norm topology coincides with the original topology, whence E is normable. □

The Kolmogorov normability criterion allows us to describe the pathology occurring for dual space of topological vector spaces beyond Banach spaces.

A.19 Proposition *Let E be a locally convex space and let*

$$E' := \{\lambda \colon E \to \mathbb{R} \mid \lambda \text{ is continuous and linear}\}$$

be its dual space. If E' is a topological vector space such that the evaluation map $\mathrm{ev} \colon E \times E' \to \mathbb{R}$, $(x,\lambda) \mapsto \lambda(x)$ *is continuous, then E is normable.*

Proof Assume that E' is a topological vector space such that ev is continuous. Then there are 0-neighbourhoods $U \Subset E$ and $V \Subset E'$ such that $\mathrm{ev}(U \times V) \subseteq [-1,1]$. Since V is absorbent by Lemma A.7 this implies that every continuous linear functional is bounded on U. Now Theorem 3.18 of Rudin (1991) yields that U is already bounded. Shrinking U, we may assume that U is convex and bounded. Hence Kolmogorov's criterion, Theorem A.18, shows that E must be normable. □

Recall from Exercise 1.4.1 that if E is normable, the evaluation map on the dual space is indeed continuous with respect to the operator norm on the dual space.

Exercises

A.2.1 Let E be a vector space and p a seminorm on E.

(a) Show that $\ker p := \{x \in E \mid p(x) = 0\}$ is a vector subspace of E.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(b) Prove that $\|q(x)\| := \inf_{y \in \ker p} p(x+y)$ defines a norm on the quotient, where q is the (surjective!) quotient map $q\colon E \to E/\ker p$.

A.2.2 Let E be a locally convex space whose topology $\mathcal{T}$ is generated by a family $\mathcal{P}$ of seminorms. Show that if q is a continuous seminorm on E, then the topology generated by $\mathcal{P} \cup \{\mathrm{q}\}$ is equal to $\mathcal{T}$.

A.2.3 Show that every locally convex space admits a fundamental system of seminorms.
Hint: Show that $\max\{p,q\}$ is a continuous seminorm and use the previous exercise.

A.3 Subspaces of Locally Convex Spaces

In this section we recall some material on subspaces of locally convex spaces.

A.20 Definition A vector subspace $F \subseteq E$ of a locally convex space is called *complemented* if there exists a locally convex space X such that $E \cong E \times X$ (isomorphic as locally convex spaces).

A.21 Lemma *A subspace $F \subseteq E$ is complemented if and only if there exists a continuous linear map $\pi\colon E \to E$ with $\pi(E) = F$ and $\pi \circ \pi = \pi$. Further, a complemented subspace is always closed. We call π a* continuous projection.

Proof Let F be complemented with isomorphism $\varphi\colon E \to F \times X$. Then $\pi := \varphi^{-1} \circ p_1 \circ \varphi$ with $p_1\colon F \times X \to F \times X$, $(f,x) \mapsto (f,0)$ is a continuous projection.

Conversely, let $\pi\colon E \to E$ be a continuous projection with $\pi(E) = F$. Then $X := \ker \pi$ is a closed subspace of E and $F \times X \to E$, $(f,x) \mapsto f + x$ is a continuous linear map with continuous inverse $e \mapsto (\pi(e), e - \pi(e))$.

If F is complemented, we have an associated continuous projection and see that $F = \ker(\mathrm{id}_E - \pi)$ is closed. □

A.22 Example Finite-dimensional subspaces of locally convex spaces are always complemented (Rudin, 1991, Lemma 4.21): Thus all subspaces of a finite-dimensional locally convex space are complemented. More generally, every closed subspace of a Hilbert space is complemented (Rudin, 1991, Theorem 12.4).

Note, however, that complemented subspaces (e.g. of Banach spaces) may be rare. Indeed one can prove that a Banach space for which every closed subspace is complemented must already be a Hilbert space (see Lindenstrauss and

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Tzafriri, 1971). Moreover, there are examples of infinite-dimensional Banach spaces whose only complemented subspaces are finite dimensional.

A.23 Example Consider the Banach space c_0 of all (real) sequences converging to 0 as a subspace of the Banach space ℓ^∞ of all bounded real sequences, with the norm

$$\|(x_1, x_2, \ldots)\|_\infty = \sup_{n\in\mathbb{N}} |x_n|.$$

Then c_0 is not complemented in ℓ^∞. The proof is, however, more involved, and so we defer to Werner (2000, Satz IV.6.5).

Exercises

A.3.1 Let $(E_i)_{i\in I}$ be a family of locally convex spaces. Prove that $\prod_{i\in I} E_i := \{(x_i)_I \mid x_i \in E_i\}$ with componentwise addition and scalar multiplication and endowed with the product topology is a locally convex space.

A.3.2 Show that $F \subseteq E$ is complemented if and only if the projection $q\colon E \to E/F$ has a continuous linear right inverse $\sigma\colon E/F \to E$ (i.e. $q \circ \sigma = \mathrm{id}_{E/F}$).

A.4 On Smooth Bump Functions

In finite-dimensional differential geometry, one uses commonly local-to-global arguments employing smooth bump functions (also sometimes called cut-off functions) and partitions of unity. This strategy fails, in general, due to a lack of bump functions. We briefly discuss the problem and refer to Kriegl and Michor (1997, Chapter III) for more information.

A.24 Definition For a map $f\colon V \to F$ with $V \subseteq E$ and E, F locally convex spaces the *carrier* of f is the set $\mathrm{carr}(f) := \{x \in V \mid f(x) \neq 0\}$. As usual the *support* of f is defined to be the closure of the carrier.

A.25 Definition Let E be a locally convex space, $x \in E$ and $U \subseteq E$ an x-neighbourhood.

(a) A C^k-map $f\colon E \to [0,\infty[$ for $k \in \mathbb{N}_0 \cup \{\infty\}$ is a *C^k-bump function* with carrier in U if $\mathrm{carr}(f) \subseteq U$ and $f(x) = 1$.

(b) If $(U_i)_{i\in I}$ is an open cover of E, we say that a family $(f_i\colon E \to [0,\infty[)_{i\in I}$ is a *C^k-partition of unity* (subordinate to the cover) if every f_i is C^k with carrier in U_i and $\sum_i f_i(x) = 1$, for all $x \in E$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

A.26 Definition Let E be a locally convex space and $k \in \mathbb{N}_0 \cup \{\infty\}$. We say that E is

(a) *C^k-regular*, if for any neighbourhood U of a point x there exists a C^k-bump function with carrier in U and $f(x) = 1$.
(b) *C^k-paracompact*,[1] if every open cover admits a C^k-partition of unity.

Obviously similar definitions make sense if we consider instead of a locally convex space a manifold modelled on locally convex spaces. While the existence of partitions of unity hinges on topological properties of the manifold (paracompactness!), the smoothness of these partitions depends only on the availability of bump functions on the model space.

A.27 (Typical local-to-global argument) Let M be a manifold which admits smooth partitions of unity subordinate to open covers. Assume we have an object defined for every chart (U,φ) in an atlas $\mathcal{A}$, and to illustrate this we choose smooth Riemannian metrics on $TU \cong U \times H$, that is, $g_U : U \times H \times H \to \mathbb{R}$, $(u,h,k) \mapsto \langle h,k\rangle_u$. Using the chart we transport it back to the manifold, that is, $g_\varphi : TU \oplus TU \to \mathbb{R}$, $(v,w) \mapsto \langle T\varphi(v), T\varphi(w)\rangle_{\varphi(\pi(v))}$. Now choose a smooth partition of unity p_φ subordinate to the open covering $(U,\varphi)_{\varphi\in\mathcal{A}}$. Then

$$g : TM \oplus TM \to \mathbb{R}, \quad (v,w) \mapsto \sum_{\varphi\in\mathcal{A}} p_\varphi(\pi_M(v)) g_\varphi(v,w)$$

is a Riemannian metric on M. Note that if $g_\varphi(v,w)$ is not defined, $p_\varphi(\pi_M(v))$ is zero so the definition makes sense.

Recall (e.g. from Hirsch, 1994, Section 2.2) that every finite-dimensional space is C^∞-regular. It is also C^∞-paracompact, as a result by Toruńczyk (see Kriegl and Michor, 1997, Corollary 16.16) shows that every Hilbert space is C^∞-paracompact. We shall now recall some results about C^k-regularity of locally convex (and, in particular, Banach) spaces.

A.28 Proposition (Bonic and Frampton, 1966) *A locally convex space is C^k-regular if and only if the topology is initial with respect to the functions $C^k(E,\mathbb{R})$.*

Proof The initial topology with respect to the C^k-functions is generated by the subbase

$$f^{-1}(]a,b[), \qquad f \in C^k(E,\mathbb{R}), a,b \in \mathbb{R} \cup \{\pm\infty\}.$$

[1] C^0-paracompact is equivalent to the usual notion of paracompactness due to Engelking (1989, Theorem 5.1.9).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Hence it is clear that if E is C^k-regular the topology is initial with respect to the C^k functions. For the converse, consider $x \in U$, where U is open in the initial topology. Then we find $a_1, \dots, a_n, b_1, \dots, b_n$ and $f_1, \dots, f_n \in C^k(E, \mathbb{R})$ for $n \in \mathbb{N}$ such that $x \in \bigcap_{1 \le i \le n} f_i^{-1}(]a_i, b_i[) \subseteq U$. Adjusting choices of the f_i, we may assume without loss of generality that $a_i = -\varepsilon_i$ and $b_i = \varepsilon_i$ for some $\varepsilon_i > 0$. Since $\mathbb{R}$ is C^∞-regular, we pick $h \in C^\infty(\mathbb{R}, \mathbb{R})$ with $h(0) = 1$ and $h(t) = 0$, for all $|t| \ge 1$. Then $f\colon E \to \mathbb{R}$, $y \mapsto \prod_{1 \le i \le n} h(f_i(x)/\varepsilon_i)$ is a bump function with carrier in U. □

For a Banach space the existence of C^k-bump functions is tied to differentiability of the norm.

A.29 Definition Let $(E, \|\cdot\|)$ be a normed space. The norm $\|\cdot\|\colon E \to \mathbb{R}$ is *rough* if there exists an $\varepsilon > 0$ such that for every $x \in E$ with $\|x\| = 1$, there exists $v \in E$ with $\|v\| = 1$ and

$$\limsup_{t \searrow 0} \frac{\|x + tv\| + \|x - tv\| - 2}{t} \ge \varepsilon.$$

If a Banach space is C^1-regular then it does not admit a rough norm (see Kriegl and Michor, 1997, 14.11).

One can prove that the Banach spaces $(C([0,1], \mathbb{R}), \|\cdot\|_\infty)$ (i.e. the continuous functions with the compact open topology) and $(\ell_1, \|\cdot\|_1)$ (see Kriegl and Michor, 1997, 13.11 and 13.12) have rough norms, whence they are not even C^1-regular. On the other hand, since nuclear Fréchet spaces are C^∞-regular, the space $C^\infty([0,1], \mathbb{R})$ is C^∞-regular. Similar statements then hold for spaces of smooth sections into bundles. Again we refer to Kriegl and Michor (1997).

A.5 Inverse Function Theorem beyond Banach Spaces

Before we conclude this appendix on locally convex spaces, let us briefly discuss (the lack of) an important tool from calculus which is driving many basic results in (finite-dimensional) differential geometry. Many basic existence results and constructions in finite-dimensional differential geometry are more or less direct consequences of the inverse function theorem, the constant rank theorem and its cousin the implicit function theorem. Note that the inverse function and the implicit function theorems still hold in Banach spaces (Lang, 1999, I, §5), but the constant rank theorem is already more delicate; see Margalef-Roig and Domínguez (1992). Beyond Banach spaces, the situation breaks down as the following example, due to Hamilton (1982, I. 5.5.1), shows (see also 1.55).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

A.30 Example Consider the Fréchet space $C^\infty([-1,1],\mathbb{R})$ of all smooth functions from $[-1,1]$ to $\mathbb{R}$,[2] together with the differential operator

$$P\colon C^\infty([-1,1],\mathbb{R}) \to C^\infty([-1,1],\mathbb{R}), \quad P(c)(x) \coloneqq c(x) - xc(x)c'(x).$$

A computation shows that P is C^∞ with derivative $dP(c,g)(x) = g(x) - xg(x)c'(x) - xc(x)g'(x)$, that is, for $c \equiv 0$ the derivative is the identity. However, the image of P is no 0-neighbourhood in $C^\infty([-1,1],\mathbb{R})$ as it does not contain any of the functions $g_n(x) \coloneqq \frac{1}{n} + \frac{x^n}{n!}$ for $n \in \mathbb{N}$ (but $\lim_{n\to\infty} g_n = 0$ in $C^\infty([-1,1],\mathbb{R})$). In conclusion, the inverse function theorem does not hold for P.

To give a more geometric example, the exponential map of a Lie group (modelled e.g. on a Fréchet space) might fail to even be a local diffeomorphism around the identity. For example, this happens for diffeomorphism groups; see Example 3.42.

A.31 (How to recover an inverse function theorem) The calculi discussed so far are too weak to provide an inverse function theorem on their own. If one has more information (such as metric estimates in the Fréchet setting) there are inverse function theorems that can apply in more general situations. The most famous one is certainly the Nash–Moser inverse function theorem (Hamilton, 1982) which works with so-called tame maps on tame spaces. Further generalised theorems are available in the framework of Müller's bounded geometry (Müller, 2008) and Glöckner's inverse function theorems; see Glöckner (2006b, 2007) and the references therein. To keep the exposition short we do not provide details here. Note, though, that the generalisations mentioned require specific settings or certain estimates which are often hard to check in applications.

A consequence of the lack of an inverse function theorem is that in infinite-dimensional differential geometry one needs to be careful when considering the notions of immersion and submersion (see §1.7). Further, there is no general solution theory for ordinary differential equations beyond Banach spaces (even for linear differential equations!).

Exercises

A.5.1 Fill in the missing details for Example A.30: Show that the differential operator P is differentiable and compute its derivative. For $g_n(x) =$

[2] The Fréchet space structure is given pointwise operations with the compact open C^∞-topology. It is defined via the metric $d(f,g) \coloneqq \sum_{i=0}^\infty \frac{\|f-g\|_i}{2^i}$, where $\|f\|_i$ is the supremum norm of the ith derivative.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$\frac{1}{n} + \frac{x^n}{n!}$ show that $g_n \to 0$ in the compact open C^∞-topology. Let f be a smooth map on $[-1,1]$. Develop f and $P(f)$ into their Taylor series at 0. Then show that $P(f) = g_n$ is impossible.

A.6 Differential Equations beyond Banach Spaces

Here we exhibit several examples of ill-posed differential and integral equation on locally convex spaces.

A.32 Example (No solutions in incomplete spaces (Dahmen)) Consider the mapping

$$h\colon [0,1] \to C^\infty([1,2],\mathbb{R}), \quad t \mapsto (x \mapsto x^t).$$

We apply the exponential law to h. Note, however, that Theorem 2.12 is not quite sufficient as $[0,1]$ and $[1,2]$ are manifolds with boundary, whence we refer to the more general exponential law (Alzaareer and Schmeding, 2015, Theorem A). Now h is smooth if and only if $h^\wedge\colon [0,1]\times[1,2] \to \mathbb{R}$, $(t,x) \mapsto x^t = \exp(\ln(x)t)$ is smooth, that is, $h^\wedge$ is smooth on the interior of the square $[0,1]\times[1,2]$ such that the derivatives extend continuously onto the boundary. This is a trivial calculation. Since the derivative of h corresponds via the exponential law to the partial derivative of $h^\wedge$ (see Lemma 2.10), we see that $dh(t;y)(x) = (d_1 h^\wedge(t,\cdot;y))^\vee(x) = y\ln(x)x^t$. Let us define now two subspaces of $C^\infty([1,2],\mathbb{R})$ as the locally convex spaces generated by the image of h and $h' = dh(\cdot;1)$:

$$E := \operatorname{span}\{h(t) \mid t \in [0,1]\}, \qquad F := \operatorname{span}\{h'(t) \mid t \in [0,1]\}.$$

By construction $h(t) \notin F$ and $h'(t) \notin E$. This entails that $h|^E\colon [0,1] \to E \subseteq C^\infty([1,2],\mathbb{R})$ is not differentiable and shows that sequential closedness is indispensable in Lemma 1.25. As a consequence, both subspaces are not closed and, in particular, not (Mackey) complete. We see that the (trivial) differential equation $\gamma'(t) = h'(t)$ or equivalently the (weak) integral equation $\gamma(t) = \int_0^1 h'(t)\mathrm{d}t$ does not admit a solution in F.

It should not come as a surprise that in the absence of suitable completeness properties differential and integral equations may be ill posed. However, even in complete spaces relative benign (e.g. linear), differential equations do not admit solutions.

A.33 Example (Hamilton, 1982, I.5.6.1) Consider the Fréchet space $C^\infty([0,1],\mathbb{R})$ of smooth functions endowed with the compact open C^∞-topology.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Recall that the differential operator $D\colon C^\infty([0,1],\mathbb{R}) \to C^\infty([0,1],\mathbb{R})$, $f \mapsto f' = df(\cdot;1)$ is continuous and linear. Consider a solution $f\colon]-\varepsilon,\varepsilon[\to C^\infty([0,1],\mathbb{R})$ of the linear differential equation

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} f(t) &= D(f) = f', \\ f(0) &= f_0. \end{cases} \tag{A.2}$$

Apply the exponential law (Alzaareer and Schmeding, 2015, Theorem A) to the smooth map f. We obtain a smooth map $f^\wedge\colon]\varepsilon,\varepsilon[\times[0,1] \to \mathbb{R}$, $(t,x) \mapsto f(t)(x)$ such that (A.2) is equivalent to the partial differential equation

$$\partial_t f^\wedge(t,x) = \partial_x f^\wedge(t,x), \quad f(0,x) = f_0(x). \tag{A.3}$$

By the Whitney extension theorem (Whitney, 1934) there is a (non-unique) extension $F_0 \in C^\infty([0,2],\mathbb{R})$ of f_0 and it is easy to see that the function $f(t,x) := F_0(x+t), t,x \in [0,1]$ solves (A.3). Since the extension F_0 is non-unique, the solution to (A.2) is non-unique (albeit we study a linear ordinary differential equation with smooth right-hand side!).

A related example is given by the heat equation on the circle.

A.34 Example (Heat equation on $\mathbb{S}^1$ (Milnor, 1982, Example 6.3)) The heat equation on the circle $\mathbb{S}^1$ is given by

$$\partial_t f(t,\theta) = \partial_\theta^2 f(t,\theta),$$

where ∂_θ is the derivative on $\mathbb{S}^1$. Again the derivative induces a continuous linear derivative operator $D\colon C^\infty(\mathbb{S}^1,\mathbb{R}) \to C^\infty(\mathbb{S}^1,\mathbb{R})$, $f \mapsto \frac{\mathrm{d}}{\mathrm{d}\theta} f$, whence the heat equation can be understood as an ordinary differential equation on $C^\infty(\mathbb{S}^1,\mathbb{R})$. We do not go into the details concerning solutions of this equation. However, the reader may want to refer to Chapter 7 for examples of partial differential equations which are treated using similar techniques for ordinary differential equations on infinite-dimensional manifolds.

The following examples are due to Milnor (see Milnor, 1982, Examples 6.1 and 6.2).

A.35 Example (Too many solutions) Let $\mathbb{R}^\mathbb{N}$ be the Fréchet space of real-valued sequences (with the topology induced by identifying $\mathbb{R}^\mathbb{N} \cong \prod_{n\in\mathbb{N}} \mathbb{R}$) and define the *left shift*

$$\Lambda\colon \mathbb{R}^\mathbb{N} \to \mathbb{R}^\mathbb{N}, \quad (x_1,x_2,x_3,\ldots) \mapsto (x_2,x_3,\ldots).$$

Then Λ is continuous and linear and the differential equation $\mathbf{y}'(t) = \Lambda(\mathbf{y})(t)$ reduces to the system of equations $y_i'(t) = y_{i+1}(t), i \in \mathbb{N}$. For every initial value this system has infinitely many solutions.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

To see this consider the initial condition $\mathbf{y}(0) = \mathbf{0} \in \mathbb{R}^{\mathbb{N}}$. If $\varphi \in C^\infty(\mathbb{R},\mathbb{R})$ vanishes in a neighbourhood of 0, then we set $y_1(t) = \varphi(t)$ and $y_n(t) = \frac{\mathrm{d}^{n-1}}{\mathrm{d}t^{n-1}}\varphi(t)$ for $n \geq 2$ to obtain a solution $\mathbf{y}(t) = (y_1(t), y_2(t), \ldots)$ of the equation with $\mathbf{y}(0) = \mathbf{0}$.

A.36 Example (No solutions) The space $E := \{(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \text{almost all } x_n = 0\}$ is locally convex with respect to the box topology (i.e. the topology whose basis is given by the sets $\prod_{n\in\mathbb{N}} U_n, U_n \Subset \mathbb{R}$; see Exercise A.6.2). Define the *right shift*

$$R\colon E \to E, \quad (x_1, x_2, x_3, \ldots) \mapsto (0, x_1, x_2, x_3, \ldots).$$

Then it is not hard to see that R is continuous linear and we consider the initial value problem

$$\begin{cases} \mathbf{y}' &= R(\mathbf{y})(t), \\ \mathbf{y}(0) &= (1,0,0\ldots). \end{cases}$$

Note that for a prospective solution $\mathbf{y}(t) = (y_1(t), y_2(t), \ldots)$, the differential equation yields $y_1'(t) = 0$, whence $y_1(t) \equiv 1$ by the initial condition. Then $y_2'(t) = y_1(t) = 1$. Integrating, we see that inductively y is a solution if $y_i(t) = \frac{t^{i-1}}{(i-1)!}$. However, for $t \neq 0$ this sequence has infinitely many terms not equal to 0 and thus does not exist in E, that is, the initial value problem (given by a linear differential equation with smooth right-hand side!) does not have any solution in E.

Exercises

A.6.1 Review §2.2 and Theorem 2.12 to work out the details of the identification of the Partial Differential Equations heat equation with the Ordinary Differential Equations on $C^\infty(\mathbb{S}^1,\mathbb{R})$ in Example A.34.

A.6.2 Consider the space $E := \{(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \text{almost all } x_n = 0\}$ with the box topology (i.e. the topology induced by $E \cap \prod_{n\in\mathbb{N}} U_i$, where $U_i \Subset \mathbb{R}$). Show that:

(a) E is a locally convex space and the box topology is properly finer than the subspace topology induced by $E \subseteq \mathbb{R}^{\mathbb{N}}$. Then show that R is continuous.

(b) Every base of 0-neighbourhoods in E is necessarily uncountable, so E cannot be a metrisable space.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

A.7 Another Approach to Calculus: Convenient Calculus

Convenient calculus was introduced by Frölicher and Kriegl (1988) (see Kriegl and Michor (1997) for an introduction). To understand the basic idea, we recall a theorem by Boman.

A.37 Theorem (Boman's theorem (Boman, 1967)) *A map $f\colon \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$ is smooth if and only if for each smooth curve $c\colon \mathbb{R} \to \mathbb{R}^d$ the curve $f \circ c$ is smooth.*[3]

Now smoothness of curves $c\colon \mathbb{R} \to E$ with values in a locally convex space E is canonically defined (e.g. via Definition 1.3). Hence smooth curves can be used to define (conveniently) smooth maps on locally convex spaces which are Mackey complete.[4] Following the usual lingo of convenient calculus, a locally convex space which is Mackey complete is called *convenient vector space*.

A.38 Definition Let E, F be convenient vector spaces.

(a) Write $c^\infty E$ for E endowed with the final topology with respect to all smooth curves $\mathbb{R} \to E$. We call this topology the *c^∞-topology* and its open sets *c^∞-open*.
(b) Let $U \subseteq E$ be c^∞-open, $f\colon U \to F$ a map. Then f is *convenient smooth* or C^∞_{conv} if $f \circ c\colon \mathbb{R} \to F$ is a smooth curve for every smooth curve $c\colon \mathbb{R} \to U$.

Obviously, the chain rule holds for conveniently smooth mappings and one can define and study derivatives. Once again manifolds, tangent spaces and so on make sense. Further, Boman's theorem asserts that between finite-dimensional spaces convenient smooth coincides with Fréchet smooth (more on this in A.41).

A.39 Remark (Bornology vs. topology) By now, the reader will have wondered why one defines C^∞_{conv}-maps on c^∞-open subsets instead of using the native topology on E. The reason for this is that differentiability of curves into a locally convex space does not depend on the topology of E, but rather on the bounded sets, the *bornology* of E. One can show that smoothness of curves is a bornological concept, and this is captured by the c^∞-topology (which is finer than the native topology but induces the same bornology).

The c^∞-topology is somewhat delicate to handle. For example, $c^\infty(E \times F) \neq c^\infty E \times c^\infty F$ (in general) and $c^\infty E$ will not be a topological vector space. However, it can be shown that for a Fréchet space the c^∞-topology coincides with the Fréchet space topology; see Kriegl and Michor (1997, Section I.4).

[3] We emphasise here that smoothness can be tested against smooth curves, while this becomes false for finite orders of differentiability.
[4] See Definition 1.12. Mackey completeness is weaker than sequential completeness.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

By definition, a conveniently smooth map is continuous with respect to the c^∞-topology on E. However, as the c^∞-topology is finer than the native locally convex topology, conveniently smooth maps may fail to be continuous with respect to the native topologies. To stress it once more: **Differentiability in the convenient calculus is *not built on top of continuity***. Having now defined two notions of calculus we will clarify their relation in the next section and discuss some of their properties.

Bastiani versus Convenient Calculus

We have already seen in the last sections that both the convenient and the Bastiani calculus yield the well-known concept of (Fréchet) smooth maps between finite-dimensional vector spaces. To clarify the relation between Bastiani and convenient calculus, observe that the definition of smooth curves into a locally convex space coincides in both calculi. The Bastiani chain rule, Proposition 1.23, yields the following.

A.40 Lemma *Let E, F be convenient vector spaces, $U \subseteq E$ and $f\colon E \supseteq U \to F$ a C^∞-map. Then f is C^∞_{conv}.*

Thus (completeness properties aside) Bastiani smoothness is the stronger and more restrictive concept, which enforces continuity with respect to the native topologies. However, on Fréchet spaces both calculi coincide (Bertram et al., 2004, Theorem 12.4).

A.41 Let E, F be convenient spaces, $U \subseteq E$. Then the differentiability classes of a map $f\colon U \to F$ are related as follows:

$$C^\infty_{\mathrm{conv}} \underset{E\,\text{Fréchet}}{\Longleftarrow} C^\infty \underset{E,F\,\text{normed}}{\Longleftrightarrow} FC^\infty. \tag{A.4}$$

The dividing line between convenient calculus and Bastiani calculus is continuity; see Glöckner (2006a) for examples of discontinuous conveniently smooth maps. Also see Kriegl and Michor (1997, Theorem 4.11) for more information on spaces on which the concepts coincide.

One may ask oneself now if there is one calculus which is preferable over the other.

A.42 (Bastiani calculus is more convenient than convenient calculus) A major difference between Bastiani and convenient calculus is continuity. Arguably continuity with respect to the native topologies, as in the Bastiani calculus, is desirable for smooth maps. In particular, the infinite-dimensional spaces and

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

manifolds often have an intrinsic topology one would rather like to preserve instead of having to deal with the somewhat delicate c^∞-topology. This is one reason why in infinite-dimensional Lie theory, Bastiani calculus is prevalent (continuity allows one to use techniques from topological group theory, such as the local-to-global result for Lie groups: Proposition 3.45).

In addition, one can often conveniently establish Bastiani smoothness using induction arguments over the order of differentiability[5] and one can interpret this (together with the continuity) as an argument for the naturality of Bastiani calculus.

A.43 (Convenient calculus is more convenient than Bastiani calculus) Discarding continuity for smooth maps might not be as exotic after all since smoothness of curves is a bornological and not a topological property. Further, it leads to a convenient category of spaces with smooth maps.

To explain this, consider sets X, Y, Z and denote by Z^X the set of all maps from X to Z. Then $f \in Z^{X\times Y}$ induces a map $f^\vee \in (Z^Y)^X$ via $f^\vee(x)(y) = f(x,y)$. The resulting bijection $Z^{X\times Y} \cong (Z^Y)^X$ is known as the exponential law.

Exponential Law for Convenient Smooth Maps *Let E, F, G be locally convex vector spaces, $U \subseteq E, V \subseteq F$ c^∞-open. Then the spaces of convenient smooth maps admit a locally convex topology such that there is a linear convenient smooth diffeomorphism*

$$C^\infty_{\mathrm{conv}}(U \times V, G) \cong C^\infty_{\mathrm{conv}}(U, C^\infty_{\mathrm{conv}}(V,G)), \quad f \mapsto f^\vee.$$

In particular, $f^\vee : U \to C^\infty_{\mathrm{conv}}(V,G)$ is C^∞_{conv} if and only if $f : U \times V \to G$ is C^∞_{conv}.

The exponential law is an immensely important tool, simplifying many proofs (e.g. that the diffeomorphism group is a Lie group). It also establishes cartesian closedness of the category of convenient vector spaces with convenient smooth maps.[6] Note that the exponential law and cartesian closedness do not hold in the Bastiani setting.[7] We remark, though, that cartesian closedness in the convenient setting is a statement about the category of convenient vector spaces. The result does not carry over to the category of manifolds modelled on convenient vector spaces which turns out to be *not cartesian closed*. To

[5] Here one should mention that a similar but somewhat more involved notion of k-times Lipschitz differentiable mappings exists in the convenient setting.

[6] The 'convenient' in convenient calculus references Steenrod's 'A convenient category of topological spaces' Steenrod (1967) in which a cartesian closed category of topological spaces is built.

[7] Albeit rudiments of an exponential law exist in the Bastiani setting as §2.2 shows. See also Alzaareer and Schmeding (2015) for a stronger version.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

obtain a cartesian closed category of manifolds it seems to be unavoidable to pass to even more general concepts such as diffeological spaces (see Iglesias-Zemmour, 2013).

Summing up, there is no clear-cut answer to the question of which calculus is preferable. It will depend on the application or use one has in mind whether one goes with the stronger notion of Bastiani calculus or the weaker convenient calculus (which often has more convenient tools).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Appendix B
Basic Ideas from Topology

We are assuming that the reader is familiar with the basic concepts of topology, and we review some definitions and results used for the reader's convenience. All of the concepts in this appendix are covered in standard textbooks on topology, for example, Engelking (1989). We mention again that we write $U \Subset X$ if U is an open subset of a topological space X.

B.1 Initial and Final Topologies

Let X be a set and $(f_i)_{i\in I}$ be a family of mappings $f_i : X \to Y_i$ to topological spaces Y_i. The coarsest topology on X making each f_i continuous is called the *initial topology on X with respect to the mappings $(f_i)_{i\in I}$*. Dually, if $\{g_i : Y_i \to X\}_{i\in I}$ is a family of mappings from topological spaces to X there is a finest topology making all g_i continuous. This topology is called the *final topology* with respect to the g_i.

B.1 Remark Note that the sets $\bigcap_{i\in F} f_i^{-1}(U_i)$, where $F \subseteq I$ is finite and $U_i \Subset Y_i$ form a basis of the initial topology (even if we restrict our choice of U_i to a basis of the topology of Y_i).

B.2 Lemma *The initial topology on X with respect to a family $(f_i)_{i\in I}$ is the unique topology which satisfies $g : Z \to X$ is continuous if and only if $f_i \circ g : Z \to Y_i$ is continuous for every $i \in I$.*

Proof Assume first that X carries the initial topology $\mathcal{I}$. Then clearly if g is continuous $f_i \circ g$ is continuous for every $i \in I$. Conversely, since $\mathcal{S} := \{f_i^{-1}(V_i) \mid V_i \Subset Y_i\}$ is a subbase for $\mathcal{I}$ by Remark B.1, we see that $f_i \circ g$ continuous implies that $g^{-1}(W)$ is open in Z for each $W \in \mathcal{S}$. Thus g is continuous. We conclude that $\mathcal{I}$ has the claimed property.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Now let $\mathcal{T}$ be another topology which satisfies the above property. Then the identity maps $\mathrm{id}\colon (X,\mathcal{I}) \to (X,\mathcal{T})$ and $\mathrm{id}\colon (X,\mathcal{T}) \to (X,\mathcal{I})$ are continuous, whence both topologies coincide. □

B.3 Example If X is a subset of a topological space Y, the *induced*, or *subspace, topology* is the initial topology with respect to the inclusion $\iota\colon X \to Y$, $x \mapsto x$.

B.4 Example We always endow the cartesian product $X := \prod_{i\in I} X_i$ of a family $(X_i)_{i\in I}$ of topological spaces with the *product topology*, which is the initial topology with respect to the family $(\mathrm{pr}_i)_{i\in I}$ of projections $\mathrm{pr}_j((x_i)_I) := x_j$.

Exercises

B.1.1 Let X be a topological space endowed with the final topology with respect to a family of mappings $f_i\colon X_i \to X$. Show that $g\colon X \to Y$ is continuous if and only if $g \circ f_i$ is continuous for all $i \in I$.

B.1.2 Let $(E_i)_{i\in I}$ be locally convex spaces and $E := \{(x_i)_{i\in I} \in \prod_{i\in I} E_i \mid \text{almost all } x_i = 0\}$. Endow E with the final topology with respect to the family of inclusions $\iota_j\colon E_j \to E$, $x \mapsto (x_i)_{i\in I}$, with $x_j = x$ and $x_i = 0$ if $i \neq j$. Show that:

(a) The resulting topology is the box topology, that is, the topology generated by the base of sets $E \cap \prod_{i\in I} U_i$, where for every $i \in I$, U_i runs through a topological base of E_i.

(b) If I is a *countable* set and $f_i\colon E_i \to F$ is a continuous linear map to a locally convex space F, then there exists a unique continuous linear map $f\colon E \to F$ with $f \circ \iota_j = f_j$. This proves, in particular, that the box topology turns E into the direct locally convex sum of the spaces E_i.
Hint: For continuity of f consider the preimage of a 0-neighbourhood V in F. Construct inductively a sequence $(V_n)_{n\in\mathbb{N}}$ of 0-neighbourhoods with $V_n + V_n \subseteq V_{n+1}$.

(c) Let I be uncountable and $E_i = \mathbb{R}$ for all $i \in I$. Show that the summation map $s\colon \{(x_i)_i \in \mathbb{R}^I \mid x_i = 0 \text{ for almost all } i\} \to \mathbb{R}$, $(x_i) \mapsto \sum_{i\in I} x_i$ is discontinuous in the box topology. Thus the box topology is properly coarser than the direct sum topology in this case.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

B.2 The Compact Open Topology

Let X, Y be (Hausdorff) topological spaces and $C(X,Y)$ the set of all continuous mappings from X to Y. We define a topology on $C(X,Y)$ by declaring a subbase consisting of the following sets:

$$\lfloor K, U \rfloor := \{f \in C(X,Y) \mid f(K) \subseteq U\}, \qquad K \subseteq X \text{ compact}, \; U \subseteq Y.$$

The resulting topology is called the *compact open topology* and we write $C(X,Y)_{\text{c.o.}}$ for the set of continuous mappings with this topology.[1]

B.5 Remark As singletons are compact, we see that the inclusion map

$$C(X,Y)_{\text{c.o.}} \to Y^X := \prod_{x \in X} Y$$

is continuous if we equip the right-hand side with the product topology. Thus $C(X,Y)_{\text{c.o.}}$ will again be Hausdorff and the evaluations $\mathrm{ev}_x \colon C(X,Y)_{\text{c.o.}} \to Y, \mathrm{ev}_x(f) := f(x)$ are continuous for every $x \in X$.

We will now show that if the target is a locally convex space the compact open topology coincides with the topology of compact convergence.

B.6 Let $(E, \{p_i\}_{i \in I})$ be a locally convex space and $K \subseteq X$ be a compact subset of a topological space. Then we define a seminorm on $C(X,E)_{\text{c.o.}}$ via

$$\|f\|_{p_i,K}(f) := \sup_{x \in K} p_i(f(x)).$$

Note that these seminorms are separating, since for $\gamma \in C(X,E)$ with $\gamma \neq 0$ we find $x \in X$ with $\gamma(x) \neq 0$ and thus $\|\gamma\|_{p_i,\{x\}} \neq 0$ for some $i \in I$. The locally convex topology generated by all seminorms $(\|\cdot\|_{p_i,K})_{i \in I,\ K \subseteq X \text{ compact}}$ is called *topology of compact convergence*.

B.7 Lemma *Let X be a topological space and E a locally convex space. Then the compact open topology coincides with the topology of compact convergence on $C(X,E)$. As a consequence, $C(X,E)_{\text{c.o.}}$ is again a locally convex space if E is locally convex.*

Proof First note that for every seminorm $\|\cdot\|_{K,p}$ for $K \subseteq X$ compact and p a continuous seminorm on E, we have

$$\{f \in C(X,E) \mid \|f\|_{K,p} < r\} = \lfloor K, \{x \in E \mid p(x) < r\} \rfloor. \tag{B.1}$$

To see that the topology of compact convergence is finer than the compact open topology, it suffices to prove that all sets $\lfloor K, U \rfloor$ with $K \subseteq X$ compact and

[1] If you prefer a video walkthrough covering (parts of) the material in this appendix, then take a look at www.youtube.com/watch?v=vGs-C9eEdJ0.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$U \Subset E$ are open in that topology. Now let $\gamma \in \lfloor K, U \rfloor$. Then $\gamma(K) \subseteq U \Subset E$ is compact and by Lemma A.2(g) there is an open 0-neighbourhood $W \Subset E$ such that $\gamma(K) + W \subseteq U$. We choose a seminorm p and $r > 0$ such that $V := \{x \in E \mid p(x) < r\} \subseteq W$. Then $\gamma + \lfloor K, V \rfloor$ is an open neighbourhood of γ in the topology of compact convergence by (B.1). Moreover, $\gamma + \lfloor K, V \rfloor$ is contained in $\lfloor K, U \rfloor$ since $\gamma(x) + \eta(x) \in \gamma(K) + V \subseteq U$ for all $\eta \in \lfloor K, V \rfloor$ and $x \in K$. This shows that $\lfloor K, U \rfloor$ is a γ-neighbourhood in the topology of compact convergence, hence open.

Conversely, let $\gamma \in C(X, E)$. Thanks to Lemma A.2(e) and (B.1), we see that the sets $\gamma + \lfloor K, \{y \in E \mid p(y) < r\} \rfloor$ form a basis of open γ-neighbourhoods in the topology of compact convergence. Thus it suffices to prove that these sets are open in the compact-open topology. For this choose $0 \in V \Subset E$ such that $V - V \subseteq \{y \in E \mid p(y) < r\}$. As $\gamma|_K$ is continuous, we find for every $x \in K$ an x-neighbourhood $K_x \Subset K$ with $\gamma(\overline{K}_x) \subseteq \gamma(x) + V$. Using compactness, we choose a finite set $F \subseteq K$ with $K = \bigcup_{x \in F} K_x$. We will now show that

$$\Omega_\gamma := \bigcap_{x \in F} \lfloor \overline{K}_x, \gamma(x) + V \rfloor \subseteq \gamma + \lfloor K, \{y \in E \mid p(y) < r\} \rfloor.$$

If $\eta \in \Omega_\gamma$, then we find for every $y \in K$ a $x \in F$ with $y \in K_x$. Thus $(\gamma - \eta)(y) \in \gamma(x) + V - (\gamma(x) + V) \subseteq V - V \subseteq U$ and thus $\gamma - \eta \in \lfloor K, \{y \in E \mid p(y) < r\} \rfloor$. Now Ω_γ is open in the compact-open topology and contains γ by choice of the K_x. Thus both topologies coincide. □

B.8 Lemma *Let A, X, Y, Z be topological spaces and $h\colon A \to X, f\colon Y \to Z$ be continuous maps. Then the* pushforward *and the* pullback *map*

$$f_*\colon C(X,Y)_{\text{c.o.}} \to C(X,Z)_{\text{c.o.}}, \quad g \mapsto f \circ g,$$
$$h^*\colon C(X,Y)_{\text{c.o.}} \to C(A,Y)_{\text{c.o.}}, \quad g \mapsto g \circ h$$

are continuous.

Proof We begin with the pushforward f_* and show that $f_*^{-1}(\lfloor K, U \rfloor)$ is open for every $K \subseteq X$ compact and $U \Subset Y$. For $\gamma \in C(X,Y)$ we see that

$$\begin{aligned} f_*(\gamma) = f \circ \gamma \in \lfloor K, U \rfloor &\Leftrightarrow f(\gamma(K)) \subseteq U \Leftrightarrow \gamma(K) \subseteq f^{-1}(U) \\ &\Leftrightarrow \gamma \in \lfloor K, f^{-1}(U) \rfloor. \end{aligned}$$

Thus $f_*^{-1}(\lfloor K, U \rfloor) = \lfloor K, f^{-1}(U) \rfloor \Subset C(X,Y)_{\text{c.o.}}$ by continuity of f.

Now for the continuity of the pullback h^*, pick $L \subseteq A$ compact and $V \Subset X$. Then for $\gamma \in C(X,Y)$, we have $h^*(\gamma) = \gamma \circ h \in \lfloor L, V \rfloor$ if and only if $\gamma(h(L)) \subseteq V$. In other words, if and only if $\gamma \in \lfloor h(L), V \rfloor$. By continuity of h, $h(L)$ is compact and thus $(f^*)^{-1}(\lfloor L, V \rfloor) = \lfloor h(L), V \rfloor \Subset C(X,Y)_{\text{c.o.}}$ and the pullback is continuous. □

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

B.9 Lemma *If X,Y,Z are topological spaces and $\mathrm{pr}_Y : Y \times Z \to Y$ and $\mathrm{pr}_Z : Y \times Z \to Z$ the canonical projections, then the map*

$$\Theta\colon C(X,Y\times Z)_{\text{c.o.}} \to C(X,Y)_{\text{c.o.}} \times C(X,Z)_{\text{c.o.}},\quad \gamma \mapsto \big((\mathrm{pr}_Y)_*(\gamma),(\mathrm{pr}_Z)_*(\gamma)\big)$$

is a homeomorphism, whence $C(X,Y\times Z)_{\text{c.o.}} \cong C(X,Y)_{\text{c.o.}} \times C(X,Z)_{\text{c.o.}}$.

Proof By Lemma B.8 the map Θ is continuous. Clearly Θ is a bijection and thus we only need to show that it takes open sets in a subbase to open sets to see that Θ is a homeomorphism. Now open rectangles $U \times V$, $U \Subset Y$, $V \Subset Z$ form a subbase of the product topology, whence Exercise B.2.2 shows that the sets $\lfloor K, U \times V \rfloor$, with $K \subseteq X$ compact, form a subbase of the topology on $C(X,Y\times Z)_{\text{c.o.}}$. Now $\Theta(\lfloor K,U\times V\rfloor) = \lfloor K,U\rfloor \times \lfloor K,V\rfloor$ is open in $C(X,Y)_{\text{c.o.}} \times C(X,Z)_{\text{c.o.}}$, and so Θ is open. □

B.10 Lemma *Let X,Y be topological spaces. If X is locally compact, the evaluation map*

$$\mathrm{ev}\colon C(X,Y)_{\text{c.o.}} \times X \to Y,\quad (f,x) \mapsto f(x)$$

is continuous.

Proof For $U \Subset Y$ we will show that $\mathrm{ev}^{-1}(U)$ is open in $C(X,Y)_{\text{c.o.}} \times X$. To this end, let $(\gamma,x) \in \mathrm{ev}^{-1}(U)$, that is, $\gamma(x) \in U$. By continuity of γ and local compactness of X there is a compact x-neighbourhood K such that $\gamma(K) \subseteq U$. Thus $\lfloor K,U\rfloor \times K$ is a (γ,x)-neighbourhood such that $\mathrm{ev}(\lfloor K,U\rfloor \times K) \subseteq U$. We see that $\mathrm{ev}^{-1}(U)$ is open and the evaluation is continuous. □

B.11 Proposition *Let X,Y,Z be topological spaces such that Y is locally compact. Then the composition map*

$$\mathrm{Comp}\colon C(X,Y)_{\text{c.o.}} \times C(Y,Z)_{\text{c.o.}} \to C(X,Z)_{\text{c.o.}},\quad (f,g) \mapsto g \circ f$$

is continuous.

Proof Let $K \subseteq X$ be compact and $U \Subset Z$. Pick (γ,η) with $\mathrm{Comp}(\gamma,\eta) \in \lfloor K,U\rfloor$. We have $\eta(\gamma(K)) \subseteq U$. By Exercise B.2.2 we can pick a compact neighbourhood L of $\gamma(K)$ in $\eta^{-1}(U)$ and set $W := L^\circ$ (interior). Then $\lfloor K,V\rfloor \times \lfloor L,U\rfloor$ is a neigbourhood of (γ,η) which is contained in $\mathrm{Comp}^{-1}(\lfloor K,U\rfloor)$ by construction. Thus the composition is continuous. □

We will now consider continuous mappings on cartesian products. If $f\colon X \times Y \to Z$ is continuous we can form for every $x \in X$ a mapping $f(x,\cdot)\colon Y \to Z$, $y \mapsto f(x,y)$. Since f is continuous every partial map $f(x,\cdot)$ is continuous. It

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

turns out that the mapping which assigns to each $x \in X$ the partial map $f(x,\cdot)$ is continuous as a mapping into the space of continuous functions.

B.12 Proposition *Let X,Y,Z be topological spaces and $f\colon X \times Y \to Z$ be continuous. Then $f^\vee\colon X \to C(Y,Z)_{\text{c.o.}}$, $f^\vee(x) \coloneqq f(x,\cdot)$ is continuous.*

Proof Consider $\lfloor K,U\rfloor \subseteq C(Y,Z)_{\text{c.o.}}$. We then compute

$$\begin{aligned}(f^\vee)^{-1}(\lfloor K,U\rfloor) &= \{x \in X \mid f(x,y) \in U \text{ for all } y \in K\}\\ &= \{x \in X \mid f(\{x\} \times K) \subseteq U\}\\ &= \{x \in X \mid \{x\} \times K \subseteq f^{-1}(U)\}.\end{aligned}$$

Now $f^{-1}(U)$ is an open neighbourhood of the compact set $\{x\} \times K$ in the cartesian product $X \times Y$. From Engelking (1989, Lemma 3.1.15) we deduce that there are $A \Subset X$ and $B \Subset Y$ such that $\{x\} \times K \subseteq A \times B \subseteq f^{-1}(U)$. This shows that $(f^\vee)^{-1}(\lfloor K,U\rfloor)$ is a neighbourhood for every x contained in it and thus an open set. Since sets of the form $\lfloor K,U\rfloor$ form a subbase of the compact open topology, $f^\vee$ is continuous. □

B.13 Proposition (Exponential law for the compact open topology) *Let X,Y,Z be topological spaces and Y be locally compact. Then a mapping $f\colon X \to C(Y,Z)_{\text{c.o.}}$ is continuous if and only if the map*

$$f^\wedge\colon X \times Y \to Z, \quad (x,y) \mapsto f(x)(y)$$

is continuous.

Proof If f is continuous, then $f^\wedge(x,y) = \mathrm{ev}(f(x),y)$ is continuous by Lemma B.10. Conversely, assume that $f^\wedge$ is continuous. Then a quick calculation shows that $f = (f^\wedge)^\vee$, whence f is continuous by Proposition B.12. □

It is worth noting that local compactness is a crucial ingredient to obtain the exponential law. Indeed one can prove that under some requirement to the topological spaces involved, the exponential law can only hold if the topological space Y is locally compact. See Engelking (1989, Exercise 3.4.A) for details.

Exercises

B.2.1 Prove that if X carries the final topology with respect to a family $g_i\colon Y_i \to X, i \in I$, then $f\colon X \to Z$ is continuous if and only if $f \circ g_i$ is continuous for every $i \in I$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

B.2.2 Let X, Y be topological spaces and $\mathcal{S}$ be a subbase for the topology on Y. Show that the sets $\lfloor K, V \rfloor$ with $K \subseteq X$ compact and $V \in \mathcal{S}$ form a subbase for the compact open topology on $C(X,Y)$.

B.2.3 Let X be a locally compact topological space and $K \subseteq X$ compact with $K \subseteq U \Subset X$. Show that there is a compact set $L \subseteq U$ whose interior contains K.

B.2.4 Let K be a compact topological space and $\Omega \Subset K \times Y$. Prove that the set

$$\Omega' := \{f \in C(K,Y) \mid \operatorname{graph}(f) \subseteq \Omega\}$$

is open in $C(K,Y)_{\text{c.o.}}$. Here $\operatorname{graph}(f) = \{(x, f(x)) \mid x \in K\} \subseteq K \times Y$.
Hint: If $f \in \Omega'$, then for every $x \in K$ there are open $U_x \Subset K, V_x \Subset Y$ with $(x, f(x)) \in U_x \times V_x \Subset \Omega$.

B.2.5 Use Proposition B.13 to give an alternative proof of the continuity of the composition map Comp from Proposition B.11.
Hint: What is $\text{Comp}^\wedge$?

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Appendix C
Canonical Manifold of Mappings

This appendix sketches the construction of a canonical manifold of mappings structure for smooth mappings between (finite-dimensional) manifolds (for details see Amiri et al., 2020, Appendix A). Before we begin, let us reconsider for a moment the locally convex space $C^\infty(M,E)$ (M a compact manifold and E a locally convex space; see §2.1). The topology and vector space structure allow us to compare two smooth maps f and g by measuring their difference $f - g$ on compact sets. As a general manifold N lacks an addition, we cannot mimic this construction (though the topology still makes sense!). On first sight it might be tempting to think that one could use the charts of N to construct charts for $C^\infty(M,N)$. However, if N does not admit an atlas with only one chart, there will be smooth mappings whose image is not contained in one chart. Thus the charts of N turn out to be not very useful. Instead one needs to find a replacement of the vector space addition to construct a way in which 'charts vary smoothly' over N.

C.1 Local Additions

In this section we first define a replacement for vector additions which allow us to construct a canonical manifold structure on spaces of mappings.

C.1 Definition Let M be a smooth manifold. A *local addition on* M is a smooth map

$$\Sigma\colon TM \supseteq U \to M,$$

defined on an open neighbourhood $U \subseteq TM$ of the *zero-section* of the tangent bundle $0_M := \{0_p \in T_pM \mid p \in M\}$ such that $\Sigma(0_p) = p$ for all $p \in M$,

$$U' := \{(\pi_M(v), \Sigma(v)) \mid v \in U\}$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

is open in $M \times M$ (where $\pi_M : TM \to M$ is the bundle projection) and the mapping $\theta := (\pi_M, \Sigma) : U \to U'$ is a C^∞-diffeomorphism.

C.2 Let G be a Lie group. Then G admits a local addition. To see this, let $\varphi : U \to V$ be a chart of G with $\mathbf{1} \in U$ and $\varphi(\mathbf{1}) = 0$. Then $\tilde{U} := T_0\varphi^{-1}(V)$ is open in $T_\mathbf{1}G$ and we define $\alpha_\mathbf{1} : \tilde{U} \to U$, $v \mapsto \varphi^{-1}(T_\mathbf{1}\varphi(v))$. Note that $\alpha_\mathbf{1}$ is a diffeomorphism. Now the tangent bundle of G is trivial, that is, $TG \cong G \times T_\mathbf{1}G$, Lemma 3.12, and we obtain an open set $\Omega := \bigcup_{g \in G} T\lambda_g(\tilde{U}) \cong G \times \tilde{U} \subseteq G \times T_\mathbf{1}G$. Define the smooth map

$$\Sigma : \Omega \to G, \quad v \mapsto \pi_G(v) \cdot \alpha_\mathbf{1}(T\lambda_{\pi_G(v)^{-1}}(v)).$$

We note that Ω is a neighbourhood of the zero section and $\Sigma(0_g) = g \cdot \alpha_\mathbf{1}(0_\mathbf{1}) = g \cdot \mathbf{1} = g$. Finally, $(\pi_G, \Sigma) : \Omega \to G \times G$ is a diffeomorphism (onto an open set) with inverse $(\pi_G, \Sigma)^{-1}(g,h) = T\lambda_g(\alpha_\mathbf{1}^{-1}(g^{-1}h))$.

C.3 Remark If (M,g) is a strong Riemannian manifold (see Chapter 4), then the Riemannian exponential map $\exp_g$ of g induces a local addition on M (Michor, 1980, Lemma 10.1).

We leave the proof of the following statement as Exercise C.1.1.

C.4 Lemma (Schmeding and Wockel, 2015, Lemma 7.5; or Michor, 1980, 10.11) *Let M be a manifold which admits a local addition Σ. Then TM admits a local addition (it is $T\Sigma \circ \kappa$, where κ is the flip of the double tangent bundle).*

With the help of a local addition we can pull to vector fields over a given function.

C.5 Let $f \in C^\infty(K,M)$ and assume that M admits a local addition $\Sigma : TM \supseteq \Omega \to M$. If $g \in C^\infty(K,M)$ such that

$$(f,g)(K) \subseteq (\pi_M, \Sigma)(\Omega) \subseteq M \times M,$$

define the mapping

$$\Phi_f(g) := (\pi_M, \Sigma)^{-1} \circ (f,g) \in C^\infty(K, TM).$$

A quick computation shows that $\pi_M \circ \Phi_f(g) = f$. The idea is that $\Phi_f(g)$ yields at every $x \in K$ a vector in $T_{f(x)}M$ which gets mapped by the local addition to $g(x)$. Thus the mapping $\Phi_f(g)$ measures the difference between f and g (similar to $f - g$ in the vector space case). Note that Φ_f takes its values in the subspace

$$C^\infty_f(K, TM) := \{h \in C^\infty(K, TM) \mid \pi_M \circ h = f\}$$

of mappings over the given map f.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

To make mappings of the form Φ_f into charts for $C^\infty(K,M)$, we first need to study spaces of the form $C^\infty_f(K,TM)$. In particular, we need these spaces to be locally convex spaces and will showcase this in the next section. The idea will be to relate these sets to spaces of sections of certain vector bundles. Note, however, that the space of functions $C^\infty_f(K,TM)$ depends heavily on the function f. In general, these spaces will *not* be isomorphic if we change the base function f.

Exercise

C.1.1 Let M be a manifold, TM its tangent bundle and $T^2M = T(TM)$ the double tangent bundle. Assume that M admits a local addition $\Sigma\colon TM \supseteq U \to M$.

(a) Show that there is a bundle isomorphism $\kappa\colon T^2M \to T^2M$ over the identity such that in local coordinates we have (up to identification) the identity $\kappa(x,v,u,w) = (x,u,v,w)$. We call κ the *(canonical) flip* of the double tangent bundle.

(b) Prove that $T\Sigma \circ \kappa$ is a local addition for TM. Explain then why $T\Sigma$ is not a local addition.

C.2 Vector Bundles and Their Sections

Recall that a *vector bundle* is a pair of manifolds E (*total space*) and M (*base (space)*) together with a surjective submersion $p\colon E \to M$ such that for every $x \in M$ the *fibre* $E_x := p^{-1}(x) \subseteq E$ is a real vector space and there is $x \in U_x \subseteq M$ and a diffeomorphism $\kappa_x\colon U_x \times F_x \to p^{-1}(U_x)$, a *bundle trivialisation* such that

- $p \circ \kappa_x(y,v) = y$, for all $y \in U_x$;
- for each $y \in U_x$, the map $\kappa_x(y,\cdot)\colon F_x \to p^{-1}(y)$ is a vector space isomorphism.

Finally, for two vector bundles $p_i\colon E_i \to M_i, i = 1,2$, a pair of smooth maps $F\colon E_1 \to E_2$ and $f\colon M_1 \to M_2$ is called *vector bundle morphism* if $p_2 \circ F = f \circ p_1$ and $F|_{p_1^{-1}(x)}^{p_2^{-1}(f(x))}$ is linear for every $x \in M_1$ (we also say that F is a bundle morphism over f).

C.6 (Sections of vector bundles) For a vector bundle $p\colon E \to M$, a smooth map $f \in C^\infty(M,E)$ is called a *smooth section* if $p \circ f = \mathrm{id}_M$. We denote the *set of all sections* by

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$\Gamma(E) := \{f \in C^\infty(M,E) \mid p \circ f = \mathrm{id}_M\}.$$

Since every fibre of a vector bundle is a vector space, pointwise addition and scalar multiplication of sections makes sense and $\Gamma(E)$ becomes a vector space. This space can be topologised as a locally convex space as follows.

C.7 Let $p\colon E \to M$ be a vector bundle. We pick a family of bundle trivialisations $\kappa_i\colon p^{-1}(U_i) \to U_i \times F_i, i \in I$ which cover E (i.e. the U_i cover M). If we denote by $\mathrm{pr}_{F_i}\colon U_\varphi \times F_i \to F_i$ the canonical projection, then the *principal part* of the section X in the trivialisation κ_i is defined as

$$X_{\kappa_i} := \mathrm{pr}_{F_i} \circ \kappa_i \circ X|_{U_i} \in C^\infty(U_i, F_i).$$

Note that this entails that with respect to the bundle trivialisation we have

$$\kappa_i \circ X = (\mathrm{id}_{U_i}, X_{\kappa_i}), \quad \text{for all } X \in \Gamma(E).$$

We declare a topology on $\Gamma(E)$ as the initial topology with respect to the map

$$\mathcal{I}_E\colon \Gamma(E) \to \prod_{i\in I} C^\infty(U_i,F_i), \quad X \mapsto (X_{\kappa_i})_{i\in I},$$

where the factors on the right-hand side carry the compact open C^∞-topology. Now by definition of the vector space operations we have $\mathcal{I}_E(X + rU) = \mathcal{I}_E(X) + r\mathcal{I}_E(U)$. Note that since the right-hand side is a locally convex space by Proposition 2.4, this also shows that $\Gamma(E)$ is a locally convex space. We shall see that this structure does not depend on the choice of trivialisations in Exercise C.2.1.

Now to model spaces of smooth mappings, we need a certain type of bundle constructed from a map and a vector bundle. While we study this situation we recall two ways to construct new vector bundles from given vector bundles.

C.8 (Pullback bundles and their sections) Let M and K be smooth manifolds. If $p\colon E \to M$ is a smooth vector bundle over M and $f\colon K \to M$ is a smooth map, then

$$f^*(E) := \bigcup_{x\in K} \{x\} \times E_{f(x)}$$

is a split submanifold of $K \times E$ (as it locally looks like $\mathrm{graph}(f) \times E_x$ inside $K \times M \times E_x$ around points in $\{x\} \times E_x$). We endow $f^*(E)$ with this submanifold structure. Together with the natural vector space structure on $\{x\} \times E_{f(x)} \cong E_{f(x)}$ and the map $p_f\colon f^*(E) \to K$, $(x,y) \mapsto x$, we obtain a vector bundle $f^*(E)$ over K, the so-called *pullback of E along f*. For each local trivialisation $\theta = (p|_{E|_U}, \theta_2)\colon E|_U \to U \times F$ of E and $W := f^{-1}(U)$, the map

$$f^*(E)|_W \to W \times F, \quad (x,y) \mapsto (x, \theta_2(y))$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

is a local trivialisation of $f^*(E)$. We endow

$$C^\infty_f(K,E) := \{\tau \in C^\infty(K,E) \mid p \circ \tau = f\}$$

with the topology induced by $C^\infty(K,E)$. With pointwise operations, $C^\infty_f(K,E)$ is a vector space and the map

$$\Psi\colon \Gamma(f^*(E)) \to C^\infty_f(K,E),\ \ \sigma \mapsto \mathrm{pr}_2 \circ \sigma$$

is a bijection with inverse $\tau \mapsto (\mathrm{id}_K, \tau)$. As $(\mathrm{pr}_2)_*\colon C^\infty(K, K\times E) \to C^\infty(K,E)$ is a continuous map and also $\tau \mapsto (\mathrm{id}_K,\tau) \in C^\infty(K,K) \times C^\infty(K,E) \cong C^\infty(K, K \times E)$ is continuous, we deduce that $C^\infty_f(K,E)$ is a locally convex topological vector space and Ψ is an isomorphism of topological vector spaces.

C.9 (Whitney sum of bundles) Let $p_i\colon E_i \to M, i = 1,2$ be vector bundles over the same base M. Then we can form the *direct product* of these vector bundles, which is the vector bundle

$$p_1 \times p_2\colon E_1 \times E_2 \to M \times M$$

over the product manifold $M\times M$. Consider the diagonal map $\Delta\colon M \to M\times M$, $m \mapsto (m,m)$. Then the *Whitney sum* of p_1 and p_2 is defined as the bundle

$$p_1 \oplus p_2\colon E_1 \oplus E_2 := \Delta^*(E_1 \times E_2) \to M.$$

Note that by construction we have as fibres $(E_1 \oplus E_2)_x = (E_1)_x \times (E_2)_x$ and if κ_i is a bundle trivialisation of E_i over a common open set $U \subseteq M$, then the restriction of $\kappa_1 \times \kappa_2$ to $(p_1 \oplus p_2)^{-1}(U)$ is a bundle trivialisation of the Whitney sum.

We just mention that there is a version of the exponential law for spaces of sections; see Amiri et al. (2020, Appendix A).

Exercises

C.2.1 Show that the mapping $\mathcal{I}_E\colon \Gamma(E) \to \prod_{i\in I} C^\infty(U_i,F)$ from C.7 has a closed image. Then prove that the topology does not depend on the choice of local trivialisations, that is, if $\mathcal{B} := \{\nu_j\}_{j\in J}$ is another family of trivialisations covering E, then the topologies induced by $\mathcal{I}_E$ and $\mathcal{I}_{E,\mathcal{B}}$ coincide.
Hint: If $\mathcal{A},\mathcal{B}$ are families of trivialisations, then $\mathcal{A}\cup\mathcal{B}$ is also such a family. Thus we may assume without loss of generality that $\mathcal{A} \subseteq \mathcal{B}$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

and it suffices to prove that the topology induced by $\mathcal{B}$ cannot be finer than the one induced by $\mathcal{A}$. To prove this, adapt Lemma 2.5.

C.2.2 Verify that the pullback bundle in C.8 forms a split submanifold of $K \times E$.
Hint: Construct submanifold charts by hand as for the graph of a function.

C.2.3 Show that for two vector bundles $p_i \colon E_i \to M$ there is a canonical isomorphism of locally convex spaces $\Gamma(E_1) \times \Gamma(E_2) \cong \Gamma(E_1 \oplus E_2)$.

C.3 Construction of the Manifold Structure

General Assumption We let K be a compact manifold and M be a smooth manifold which admits a local addition $\Sigma \colon TM \supseteq U \to M$ such that

- Σ is *normalised*, that is, $T_{0_p}(\Sigma|_{T_pM}) = \mathrm{id}_{T_pM}$ for all $p \in M$ (a manifold with local addition has a normalised local addition; see Amiri et al., 2020, Lemma A.14);
- $\theta := (\pi_M, \Sigma) \colon U \to U'$ is a diffeomorphism.

C.10 (Manifold structure on $C^\infty(K,M)$) For $f \in C^\infty(K,M)$, the locally convex space of C^∞-sections of $f^*(TM)$ can be identified with

$$C^\infty_f(K,E) = \{\tau \in C^\infty(K,TM) \mid \pi_M \circ \tau = f\},$$

with the topology induced by $C^\infty(K,TM)$. Use notation as in Definition C.1. Then $O_f := C^\infty_f(K,E) \cap C^\infty(K,U)$ is an open subset of $C^\infty_f(K,E)$,

$$O'_f := \{g \in C^\infty(K,M) \mid (f,g)(K) \subseteq U'\}$$

is an open subset[1] of $C^\infty(K,M)$ and the map $\phi_f \colon O_f \to O'_f$, $\tau \mapsto \Sigma \circ \tau$ is a homeomorphism with inverse $g \mapsto \theta^{-1} \circ (f,g)$. By the preceding, if also $h \in C^\infty(K,M)$, then $\phi_h^{-1} \circ \phi_f$ has an open domain; $\phi_h^{-1} \circ \phi_f$ is smooth there, since (by the exponential law, Theorem 2.12), we only need to observe that the map

$$(\tau, x) \mapsto (\phi_h^{-1} \circ \phi_f)(\tau)(x) = \theta^{-1}(h(x), \Sigma(\tau(x))) \tag{C.1}$$

is C^∞. Hence $C^\infty(K,M)$ has a smooth manifold structure such that each of the maps ϕ_f^{-1} is a local chart.

We prove that the manifold structure on $C^\infty(K,M)$ is canonical and thus by Lemma 2.16(b) the construction C.10 is independent of the choice of local addition.

[1] The proof is similar to Exercise B.2.4.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

C.11 Lemma *The manifold structure on $C^\infty(K,M)$ constructed in C.10 is canonical.*

Proof We first show that the evaluation map $\mathrm{ev}\colon C^\infty(K,M)\times K \to M$ is C^∞. It suffices to show that $\mathrm{ev}(\phi_f(\sigma),x)$ is C^∞ in $(\sigma,x)\in O_f\times K$ for all $f\in C^\infty(K,M)$. But

$$\mathrm{ev}(\phi_f(\sigma),x) = \Sigma(\sigma(x)) = \Sigma(\mathrm{ev}(\sigma,x)),$$

where $\mathrm{ev}\colon C^\infty_f(K,E)\times K \to f^*(TM)$, $(\sigma,x)\mapsto\sigma(x)$ is C^∞ (see Alzaareer and Schmeding, 2015, Proposition 3.20). Now let $h\colon L\to C^\infty(K,M)$ be a map, where L is a manifold. If h is C^∞, then $h^\wedge = \mathrm{ev}\circ(h\times \mathrm{id}_K)$ is C^∞. If, conversely, $h^\wedge$ is C^∞, then h is continuous as a map to $C(K,M)$ with the compact open topology (see Proposition B.12) and $h(x) = h^\wedge(x,\cdot)\in C^\infty(K,M)$ for each $x\in L$. Given $x\in L$, let $f := h(x)$. Then

$$\psi_f\colon C(K,M)\to C(K,M)\times C(K,M)\cong C(K,M\times M),\quad g\mapsto (f,g)$$

is a continuous map and $W := h^{-1}(O'_f) = (\phi_f\circ h)^{-1}(C(K,U'))$ is an open x-neighbourhood in L. Since

$$((\phi_f)^{-1}\circ h|_W)^\wedge(y,z) = (\theta^{-1}\circ(f,h(y)))(z) = \theta^{-1}(f(z),h^\wedge(y,z))$$

is C^∞, the map $\phi_f^{-1}\circ h|_W$ (and hence also $h|_W$) is C^∞ (apply Theorem 2.12 to the spaces $C^\infty(U_i,F)$ containing the principal parts of the sections). □

C.12 Let $\gamma\in C^\infty(K,M)$ and view $T_\gamma C^\infty(K,M)$ as a set of equivalence classes of smooth curves $c\colon\,]-\varepsilon,\varepsilon[\to C^\infty(K,M), c(0)=\gamma$. As the manifold structure is canonical, c is smooth if and only if $c^\wedge\colon\,]-\varepsilon,\varepsilon[\times K\to M$ is smooth. Hence for the canonical chart $\phi_\gamma\colon O_\gamma\to O'_\gamma\subseteq C^\infty(K,M)$, the map $T_0\phi_\gamma\colon C^\infty_\gamma(K,TM)\to\!\!\to T_\gamma C^\infty(K,M)$ is an isomorphism of TVS. For $x\in K$ denote by ε_x the point evaluation in x. Since Σ is normalised we obtain

$$\begin{aligned} T\varepsilon_x T\phi_f(0,\tau) &= T\varepsilon_x([t\mapsto\Sigma\circ(t\tau)]) = [t\mapsto\Sigma(t\tau(x))]\\ &= [t\mapsto\Sigma|_{T_{f(x)}M}(t\tau(x))] = T\Sigma|_{T_{f(x)}M}(\tau(x)) = \tau(x).\end{aligned}$$

Summing up, this implies that for each fibre there is a linear bijection

$$\Phi_\gamma\colon T_\gamma C^\infty(K,M)\to C^\infty_\gamma(K,TM),\quad [c]\mapsto(k\mapsto[t\mapsto c^\wedge(t,k)]). \tag{C.2}$$

We will now sketch the proof that the fibre maps (C.2) induce a bundle isomorphism

$$\Phi_M\colon TC^\infty(K,M)\to C^\infty(K,TM),\quad T_\gamma C^\infty(K,M)\ni V\mapsto\Phi_\gamma(V)$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

such that the following diagram commutes

$$\begin{array}{ccc} TC^\infty(K,M) & \xrightarrow{\quad\Phi_M\quad} & C^\infty(K,TM) \\ & {\scriptstyle \pi_{C^\infty(K,M)}} \searrow \quad \swarrow {\scriptstyle (\pi_M)_*} & \\ & C^\infty(K,M). & \end{array}$$

Sketch of Proof (See Amiri et al., 2020, Appendix A.) If $\lambda_p : T_pM \to TM$ is the inclusion and $\kappa : T^2M \to T^2M$ the canonical flip, then $\Theta : TM \oplus TM \to \pi_{TM}^{-1}(0_M) \subseteq T^2M, \Theta(v,w) = \kappa(T\lambda_{\pi(v)}(v,w))$ is a bundle isomorphism. Let $0 : M \to TM$ be the zero-section. Then Θ induces a diffeomorphism

$$\Theta_\gamma : O_\gamma \to O_{0\circ\gamma}, \quad \eta \mapsto \Theta \circ (0 \circ \gamma, \eta).$$

From the local addition Σ, we construct a local addition $\widetilde{\Sigma}$ on TM and consider the charts $\phi_{0\circ\gamma}$ on $C^\infty(K,TM)$. Then the sets $S_\gamma := T\phi_\gamma(O_\gamma \times C^\infty_\gamma(K,TM))$ form an open cover of $T(C^\infty(K,M))$ for $\gamma \in C^\infty(K,M)$. We deduce that the sets $\Phi_M(S_\gamma)$ form a cover of $C^\infty(K,TM)$ by sets which are open as $\Phi_M(S_\gamma) = (\phi_{0\circ\gamma} \circ \phi_\gamma)(O_\gamma \times C^\infty_\gamma(M,TM)) = \phi_{0\circ\gamma}(O_{0\circ\gamma})$. Hence we can check that Φ_M restricts to a C^∞-diffeomorphism on these open sets, that is,

$$\Phi_M \circ T\phi_\gamma = \phi_{0\circ\gamma} \circ \Theta_\gamma$$

for each $\gamma \in C^\infty(K,M)$ (as all other mappings in the formula are smooth diffeomorphisms). Now we can rewrite $\Phi_M(T\phi_\gamma(\sigma,\tau))$ as

$$\begin{aligned} &([t \mapsto \Sigma(\sigma(x) + t\tau(x))])_{x\in K} = ([t \mapsto (\Sigma \circ \lambda_{\gamma(x)})(\sigma(x) + t\tau(x))])_{x\in K} \\ &\quad = (T(\Sigma \circ \lambda_{\gamma(x)})(\sigma(x),\tau(x)))_{x\in K} = (\Sigma_{TM}((\kappa \circ T\lambda_{\gamma(x)})(\sigma(x),\tau(x))))_{x\in K} \\ &\quad = ((\Sigma_{TM} \circ \Theta_\gamma)(\sigma,\tau)(x))_{x\in K} = (\phi_{0\circ\gamma} \circ \Theta_\gamma)(\sigma,\tau). \end{aligned}$$

Thus Φ_M is a C^∞-diffeomorphism.

C.13 (Smooth maps into the Whitney sum over strong Riemannian manifolds) By Lemma C.4, TM admits a local addition whose product with itself yields a local addition on the product manifold. Thus $C^\infty(K,M), C^\infty(K,TM), C^\infty(K,TM\times TM)$ are canonical manifolds. Taking the Whitney sum of $(\pi_M)_*$: $C^\infty(K,TM) \to C^\infty(K,M)$ (C.12) with itself, we obtain the bundle $C^\infty(K,TM) \oplus C^\infty(K,TM)$. Our aim is to identify it with the bundle $C^\infty(K,TM \oplus TM)$. Observe that $C^\infty(K,TM) \oplus C^\infty(K,TM)$ is a split submanifold of $C^\infty(K,M) \times C^\infty(K,TM)^2$. Now the factors of the product are canonical manifolds. Thus Lemma 2.16 yields a diffeomorphism

$$C^\infty(K,M) \times C^\infty(K,TM)^2 \cong C^\infty(K, M \times (TM)^2),$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

which takes the split submanifold $C^\infty(K,TM) \oplus C^\infty(K,TM)$ to $C^\infty(K,TM \oplus TM)$. As diffeomorphisms preserve split submanifolds, see Exercise 4.4.2, $C^\infty(K,TM \oplus TM)$ must be a split submanifold of $C^\infty(K, M \times (TM)^2)$. Finally, Lemma 2.16(c) shows that $C^\infty(K,TM \oplus TM)$ is a canonical manifold diffeomorphic to $TC^\infty(K,M) \oplus TC^\infty(K,M)$.

C.14 Remark By uniqueness of canonical manifolds, $C^\infty(K,TM \oplus TM)$ from C.13 coincides with the manifold structure we could have obtained via a local addition on $TM \oplus TM$. The same proof works if we only assume that $C^\infty(K,M)$ and $C^\infty(K,TM)$ are canonical manifolds (without assuming that M has a local addition).

Exercise

C.3.1 Work out the missing details in the sketch of the proof in C.12.

C.3.2 Prove that a manifold with a local addition admits a normalised local addition.

C.4 Manifolds of Curves and the Energy of a Curve

In this appendix we consider the manifold structure on spaces of curves on a compact interval. The reason for this is that we defined for a (weak) Riemannian manifold (M,g) the energy $\mathrm{En}\colon C^\infty([0,1],M) \to \mathbb{R}$ of a curve and would like to differentiate En to find geodesics. Hence a manifold structure on the space of curves $C^\infty([0,1],M)$ is needed. Many details of the construction will be left to the reader as Exercise C.4.1. Moreover, we will not systematically introduce tangent bundles for manifolds with boundary such as $[0,1]$ (see e.g. Michor, 1980). Thus the compact open C^∞-topology needs to be defined without recourse to tangent bundles.

C.15 Let M be a (possibly infinite-dimensional) manifold. Let $c\colon [0,1] \to M$ be a smooth curve and $K \subseteq [0,1]$ compact such that $c(K) \subseteq U$, where (U,φ) is a chart of M. If $\varphi(U) \subseteq E$ for the locally convex space E, we pick a seminorm $\|\cdot\|$ on E and define a C^k-neighbourhood $N^k(c,K,(U,\varphi),\|\cdot\|,\varepsilon)$ as the set

$$\left\{ g \in C^\infty([0,1],M) \middle| g(K) \subseteq U, \sup_{0\le\ell\le k} \sup_{x\in K} \left\| \frac{\mathrm{d}^\ell}{\mathrm{d}x^k}(\varphi\circ g - \varphi\circ c)(x) \right\| < \varepsilon \right\}.$$

Then the family $N^k(c,K,(U,\varphi),\|\cdot\|,\varepsilon)$, where $c \in C^\infty([0,1],M)$, $K \subseteq [0,1]$ compact, $\|\cdot\|$ is a continuous seminorm of E and $\varepsilon > 0$ forms the base of a topology on $C^\infty([0,1],M)$ called the *compact-open C^∞-topology*. If $M = E$ is

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

a locally convex space, then $C^\infty([0,1],E)$ is also a locally convex space[2] and the exponential law, Theorem 2.12, carries over to $C^\infty([0,1],O)$, $O \subseteq E$.

C.16 Proposition *Let (M,Σ) be a manifold with local addition Σ and topologise the space $C^\infty([0,1],M)$ with the compact-open C^∞-topology. Then we have that $C^\infty([0,1],M)$ is a canonical manifold and $TC^\infty([0,1],M) \cong C^\infty([0,1],TM)$.*

C.17 Lemma *Let (M,g) be a (weak) Riemannian manifold such that M and TM admit local additions. Then the* energy

$$\mathrm{En}\colon C^\infty([0,1],M) \to \mathbb{R}, \quad c \mapsto \frac{1}{2}\int_0^1 g_{c(t)}(\dot c(t),\dot c(t))\mathrm{d}t$$

is smooth. We can express its derivative in a local chart (U,φ) (suppressing most identifications) as

$$d\mathrm{En}(c;h) = \int_0^1 \frac{1}{2} d_1 g_U(c,c'(t),c'(t);h) - d_1 g(c(t),h(t),c'(t);c'(t)) - g_U(c(t),h(t),c''(t))\mathrm{d}t,$$

where we view g as a map of three arguments, $g_U(c,a,b)$ and $h \in T_c C^\infty([0,1],M) \cong \{g \in C^\infty([0,1],TM) \mid \pi\circ g = c\}$ with $h(0) = 0_{c(0)}$ and $h(1) = 0_{c(1)}$.

Proof Since M and TM admit local additions, Exercise C.4.4 implies that both $C^\infty([0,1],M)$ and $C^\infty([0,1],TM)$ are canonical manifolds. Applying the exponential law, $C^\infty([0,1],M) \to C^\infty([0,1],TM)$, $c \mapsto \dot c = (t \mapsto T_t c(1))$ is smooth if and only if $C^\infty([0,1],M)\times[0,1] \to TM, (c,t) \mapsto T_t c(1)$ is smooth. We check this locally in a neighbourhood of a c: Pick a chart (U,φ) of M and $[a,b] \subseteq [0,1]$ with $c([a,b]) \subseteq U$. As the topology on $C^\infty([0,1],M)$ is finer than the compact-open topology, there exists a whole neighbourhood of curves g with $g([a,b]) \subseteq U$. Cover $[0,1]$ by compact intervals which c maps into a chart domain and work locally. To keep the notation simple, we will assume that $c([0,1]) \subseteq U$ or in other words, assume without loss of generality that $M \subseteq E$ for some locally convex vector space E. Thus we need to prove that $C^\infty([0,1],M)\times[0,1] \to M\times E$, $(c,t) \mapsto (c(t),c'(t)) = (\mathrm{ev}(c,t),\mathrm{ev}(c',t))$ is smooth. The evaluation map is smooth on canonical manifolds and the mapping $C^\infty([0,1],M) \to C^\infty([0,1],E)$, $c \mapsto c'$, is the restriction of the continuous linear map $C^\infty([0,1],E) \to C^\infty([0,1],E)$, $c \mapsto c'$, to the open subset $C^\infty([0,1],M) \subseteq C^\infty([0,1],E)$ (here we exploit the compact open C^∞-topology). We conclude that the mapping $C^\infty([0,1],M) \to C^\infty([0,1],TM)$,

[2] For $M = \mathbb{R}$, we have described this structure already in Example 1.6.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$c \mapsto \dot{c}$ is smooth. Since Exercise C.4.6 identifies the Whitney sums, we deduce from $(\dot{c},\dot{c}) \in TM \oplus TM$ that $C^\infty([0,1],M) \to C^\infty([0,1],\mathbb{R})$, $c \mapsto g_*(\dot{c},\dot{c})$ is smooth (as pushforwards are smooth on canonical manifolds). However, $\frac{1}{2}\int_0^1 : C^\infty([0,1],\mathbb{R}) \to \mathbb{R}$ is continuous linear, whence En can be written as a composition of smooth mappings and is thus smooth.

To compute the derivative of the energy, we work in a local chart (U,φ) of M (though we will only label g and suppress the other identifications). Then the metric g becomes a map of three arguments g_U which is bilinear in the last two. Recall that the vector component of $\dot{c}$ is c'. Now by choice of h, there is a smooth curve $q\colon]-\varepsilon,\varepsilon[\to C^\infty([0,1],M)$ such that $\frac{\partial}{\partial s}\big|_{s=0} q(s) = h$ and $q(t)(0) = c(0)$ and $q(t)(1) = c(1)$ for all t (a smooth variation; see also Definition 7.1 for the meaning of the partial derivative). Then we compute, with the help of the exponential law (Exercise C.4.3),

$$
\begin{aligned}
d\mathrm{En}(c;h) &= \frac{d}{ds}\Big|_{s=0} \mathrm{En}(q(s)) \\
&= \frac{1}{2}\int_0^1 \frac{d}{ds}\Big|_{s=0} g_U\left(q(s)(t), \frac{d}{dt}q(s)(t), \frac{d}{dt}q(s)(t)\right)\mathrm{d}t \\
&= \int_0^1 \frac{1}{2} d_1 g_U(c(t),c'(t),c'(t);h(t)) + g_U\left(c(t), \frac{d}{ds}\Big|_{s=0}\frac{d}{dt}q(s)(t), c'(t)\right)\mathrm{d}t \\
&= \int_0^1 \frac{1}{2} d_1 g_U(c(t),c'(t),c'(t);h(t)) - d_1 g_U\left(c(t), \frac{d}{ds}\Big|_{s=0} q(s)(t), c'(t); c'(t)\right) \\
&\qquad\qquad - g_U\left(c(t), \frac{d}{ds}\Big|_{s=0} q(s)(t), c''(t)\right)\mathrm{d}t \\
&= \int_0^1 \frac{1}{2} d_1 g_U(c(t),c'(t),c'(t);h(t)) - d_1 g_U(c(t),h(t),c'(t);c'(t)) \\
&\qquad\qquad - g_U(c(t),h(t),c''(t))\mathrm{d}t.
\end{aligned}
$$

In passing from the second to the third lines we used integration by parts together with the fact that $\frac{d}{ds}\big|_{s=0} q(s,t)$ vanishes at the endpoints of the interval. □

Almost all of the terms in the formula for the derivative of the energy in Lemma C.17 can be globalised to the Riemannian manifold. Derivation exploits of course that we work locally, and the second derivative of c needs to be taken (from the perspective of the Riemannian manifold M) in the fibre over $c(t)$. This already hints at the connection of this formula to the covariant derivative (which, however, was not yet needed).

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Exercises

C.4.1 Prove that the neighbourhoods defined in C.15 form the base of a topology.

C.4.2 Let E be a locally convex space. Show that the compact-open C^∞-topology turns $C^\infty([0,1],E)$ into a locally convex space. Show then that for $M = \mathbb{R}$ this topology coincides with the compact-open C^∞-topology from Example 1.6.

C.4.3 Establish a variant of the exponential law, Theorem 2.12, for manifolds of smooth mappings on $[0,1]$ (with values in open sets of locally convex spaces).

C.4.4 Generalise C.10 to prove Proposition C.16. Then proceed to show that $C^\infty([0,1],M)$ is a canonical manifold.

C.4.5 Follow the argument in C.12 to prove that $TC^\infty([0,1],M)$ can naturally be identified with $C^\infty([0,1],TM)$.

C.4.6 Adapt the argument in C.13 to establish an isomorphism

$$C^\infty([0,1],TM \oplus TM) \cong C^\infty([0,1],TM) \oplus C^\infty([0,1],TM)$$

if M and TM admit local additions.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Appendix D
Vector Fields and Their Lie Bracket

D.1 Construction

In this appendix we recall the construction of the Lie algebra of vector fields.

D.1 Example The tangent bundle $\pi_M : TM \to M$ is a vector bundle (bundle trivialisations are given by the canonical charts $T\varphi$). A smooth section of the tangent bundle is called (smooth) *vector field* and we write shorter $\mathcal{V}(M) := \Gamma(TM)$ for the vector space of all vector fields.

D.2 Example If $U \subseteq E$ in a locally convex space, we have $TU = U \times E$ and $\pi_U : U \times E \to U$, $(u,e) \mapsto u$. Thus a vector field of U can be written as $X = (X_U, X_E) \colon U \to U \times E$ and we must have $X_U = \mathrm{id}_U$. Hence a vector field on U is uniquely determined by the smooth map $X_E \in C^\infty(U,E)$.

D.3 If M is a manifold and (ϕ, U_ϕ) a manifold chart, then we have an analogue of X_E on U_ϕ for $X \in \mathcal{V}(M)$: Clearly $T\phi \circ X \circ \phi^{-1} = (\mathrm{id}_{V_\phi}, X_\phi)$ for the smooth map $X_\phi := d\phi \circ X \circ \phi^{-1} \colon V_\phi \to E$. We call X_ϕ the *local representative of* X or *the principal part of* X with respect to the chart ϕ.

For later use, consider a vector field $X \in \mathcal{V}(M)$ and a smooth function $f \colon M \to F$, where F is a locally convex space. Then we define a function $X.f \in C^\infty(M,F)$ via

$$X.f(m) := df \circ X(m) = \mathrm{pr}_2 \circ Tf \circ X(m). \tag{D.1}$$

D.4 Similar to C.7 we topologise $\mathcal{V}(M)$: Pick an atlas $\mathcal{A}$ of M whose charts we denote by $\varphi \colon U_\varphi \to V_\varphi \subseteq E_\varphi$. Then we declare the topology to be the initial topology with respect to the map

$$\kappa \colon \mathcal{V}(M) \to \prod_{\varphi \in \mathcal{A}} C^\infty(V_\varphi, E_\varphi), \quad X \mapsto (X_\varphi)_{\varphi \in X},$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

where the factors on the right-hand side carry the compact open C^∞-topology. In particular, this topology turns the vector fields into a locally convex space.

We will use the notion of integral curves and flows for vector fields, whence we recall the definition of these objects.

D.5 Let $X \in \mathcal{V}(M)$. We say a C^1-curve $c\colon [a,b] \to M$ is an *integral curve* for X if for every $t \in [a,b]$ the curve satisfies $\dot{c}(t) = X(c(t))$.

If M is a Banach manifold, it follows from the theory of ordinary differential equations, Lang (1999, IV), that for every $m \in M$ there exists an integral curve c_m of X on some open interval $J_m :=]-\varepsilon,\varepsilon[$ such that $c_m(0) = m$. Moreover, the *flow*

$$\mathrm{Fl}^X\colon \bigcup_{m\in M} \{m\} \times J_m \to M, \quad (m,t) \mapsto c_m(t)$$

defines a continuous map on some open subset of $M \times \mathbb{R}$. If M is modelled on a locally convex space, the existence of integral curves and flows is not automatic; see Appendix A.6.

D.6 Definition Let $f\colon M \to N$ be smooth. We call the vector fields $X \in \mathcal{V}(M), Y \in \mathcal{V}(N)$ *f-related* if $Y \circ f = Tf \circ X$.

D.7 Lemma *Let M be a manifold modelled on a locally convex space E with atlas $\mathcal{A}$. Let $(X_\phi)_{\phi\in\mathcal{A}}$ be a family of smooth maps $X_\phi\colon V_\phi \to E$ such that every pair X_ϕ, X_ψ is $\psi \circ \phi^{-1}$ related on $\phi(U_\psi \cap U_\phi)$. Then there is a unique vector field $X \in \mathcal{V}(M)$ whose local representatives coincide with the X_ϕ.*

Proof Define $X\colon M \to TM$, $p \mapsto T\phi^{-1}(\phi(p), X_\phi(\phi(p)))$ for $p \in U_\phi$. Since the maps X_ϕ, X_ψ are related by the change of charts on the overlap $U_\phi \cap U_\psi$, the mapping is well defined. By construction it is smooth and a vector field. □

D.8 For principal parts of vector fields X, Y on $U \subseteq E$ write $X_E.Y_E(z) := dY_E \circ X(z) := dY_E(z; X_E(z))$. Define

$$[X,Y] := X.Y - Y.X. \qquad X,Y \in C^\infty(U,E).$$

We will see in the following that the bracket of principal parts of vector fields gives rise to a Lie bracket of vector fields.

D.9 Lemma *Let $U \subseteq E$, $V \subseteq F$ be open in locally convex spaces and $f \in C^\infty(U,V)$, $X_1, X_2 \in C^\infty(U,E)$ and $Y_1, Y_2 \in C^\infty(V,F)$. Assume that X_i is f-related to Y_i for $i = 1,2$. Then $[X_1,X_2]$ is f-related to $[Y_1,Y_2]$.*

Proof Using the chain rule, (1.7) and relatedness we obtain $i=1,2 (x,v)\in U\times E$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$df(x, dX_i(x;v)) = dY_i(f(x), df(x;v)) - d^2 f(x; X_i(x), v), \tag{D.2}$$

We use this relation together with relatedness to obtain

$$\begin{aligned} df(x;[X_1,X_2](x)) &= df(x;dX_2(x,X_1(x))) - df(x;dX_1(x;X_2(x))) \\ &= dY_2(f(x);df(x;X_1(x))) - d^2 f(x;X_2(x),X_1(x)) \\ &\quad - dY_1(f(x);df(x;X_2(x))) + d^2 f(x;X_1(x),X_2(x)) \\ &= dY_2(f(x);Y_1(f(x))) - dY_1(f(x);Y_2(f(x))) \\ &= [Y_1,Y_2](f(x)), \end{aligned}$$

where the second-order terms cancel by Schwarz' theorem. □

Before we now establish the Lie algebra properties, let us recall a general definition useful for our purpose.

D.10 Definition Let $(A,\cdot)$ be an associative algebra. Then the linear mappings $L(A,A)$ form a Lie algebra under the commutator bracket $[\phi,\psi] := \phi\circ\psi - \psi\circ\phi$, Example 3.16 (where $\circ$ is the usual composition of linear maps). A mapping $\phi \in L(A,A)$ is called *derivation* of the algebra A if it satisfies the Leibniz rule

$$\phi(a\cdot b) = \phi(a)\cdot b + a\cdot\phi(b) \quad \text{for all } a,b\in A.$$

We denote by $\mathrm{der}(A)$ the *set of all derivations* of A and note that it forms a Lie subalgebra of $(L(A,A),[\cdot,\cdot])$. (As no topology is involved, this will, in general, not be a locally convex Lie algebra.)

For E a locally convex space, $U \subseteq E$ and $X \in \mathcal{V}(U)$ define the *Lie derivative*

$$\mathcal{L}_X(f) := df\circ X = df(\mathrm{id}_U, X_E) \text{ for } f\in C^\infty(U,\mathbb{R}). \tag{D.3}$$

By definition $\mathcal{L}_X(f) = X.f$ in the special case that f is real valued. The reason for the new notation and name will become apparent from the following observations (see also Definition E.9): The pointwise multiplication turns $C^\infty(U,\mathbb{R})$ into an associative algebra. Then $\mathcal{L}_X$ is linear in f. Thus

$$\mathcal{L}_X(f\cdot g) = \mathcal{L}_X(f)\cdot g + f\cdot\mathcal{L}_X(g). \tag{D.4}$$

In other words, $\mathcal{L}_X$ is a derivation of the algebra $C^\infty(U,\mathbb{R})$.

D.11 Lemma *Let $U \subseteq E$ in a locally convex space.*

(a) $\mathcal{L}_{[X,Y]} = \mathcal{L}_X\circ\mathcal{L}_Y - \mathcal{L}_Y\circ\mathcal{L}_X$.
(b) *The map $\mathcal{L}\colon C^\infty(U,E)\to \mathrm{der}(C^\infty(U,R))$, $X\mapsto\mathcal{L}_X$ is linear and injective.*

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(c) *The map* $[\cdot,\cdot]\colon C^\infty(U,E)\times C^\infty(U,E)\to C^\infty(U,E)$, $(X,Y)\mapsto [X,Y]=X.Y-Y.X$ *turns the space* $C^\infty(U,E)$ *into a Lie algebra*.

Proof (a) From (1.7) we deduce

$$\mathcal{L}_X(\mathcal{L}_Y(f)) = d^2f(x;Y(x),X(x)) + df(x;dY(x;X(x))).$$

Also using the formula for X and Y interchanged, we see that the second-order terms cancel by Schwarz' theorem and thus

$$[\mathcal{L}_X,\mathcal{L}_Y](f)(x) = \mathcal{L}_{[X,Y]}(f)(x).$$

(b) $\mathcal{L}_X$ is linear in X as $df(x;\cdot)$ is. Thus it suffices to prove that the kernel of $\mathcal{L}$ is trivial. Let $X\in C^\infty(U,E)$ be a map with $X(x)\neq 0$ for some $x\in U$. By the Hahn–Banach theorem, 1.7, we find $\lambda\in E'$ with $\lambda(X(x))\neq 0$. Then $\mathcal{L}_X(\lambda)(x) = d\lambda(x,X(x)) = \lambda(X(x))\neq 0$ and thus $\mathcal{L}_X\neq 0$.

(c) Clearly $[\cdot,\cdot]$ is bilinear, whence $(C^\infty(U,E),[\cdot,\cdot])$ is an algebra. Now $[X,X]=X.X-X.X=0$. Recall that in the Jacobi identity, the entries of the iterated Lie bracket are cyclically permuted. We write shorter $\sum_{\text{cycl}}[X,[Y,Z]]$ for this and thus have to check that this expression vanishes for all $X,Y,Z\in C^\infty(U,E)$. However,

$$\mathcal{L}\left(\sum_{\text{cycl}}[X,[Y,Z]]\right) = \sum_{\text{cycl}}[\mathcal{L}_X,[\mathcal{L}_Y,\mathcal{L}_Z]] = 0,$$

where we have used linearity of $\mathcal{L}$, (a), (b) and the fact that the derivations form a Lie algebra. Since $\mathcal{L}$ is injective by (b), we see that the Jacobi identity holds. □

Finally, we show that the Lie bracket of vector fields is continuous if the space E is finite dimensional.

D.12 Lemma *Let E be a finite-dimensional space and $U\subseteq E$. Then the Lie bracket*

$$[\cdot,\cdot]\colon C^\infty(U,E)\times C^\infty(U,E)\to C^\infty(U,E)$$

is continuous. Hence $(C^\infty(U,E),[\cdot,\cdot])$ is a locally convex Lie algebra.

Proof Note that $C^\infty(U,E)$ is a locally convex space with respect to the compact open C^∞-topology, Proposition 2.4. To establish continuity of the Lie bracket, we deduce from Lemma 2.10 that it suffices to establish continuity of the adjoint map

$$p\colon C^\infty(U,E)\times C^\infty(U,E)\times U\to E,\quad (X,Y,u)\mapsto dY(u;X(u)).$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Recall that the compact-open C^∞-topology is initial with respect to the mappings $d^k\colon C^\infty(U,E) \to C(U \times E^k, E)_{\mathrm{c.o.}}$, $f \mapsto d^k f$. Hence the map $d\colon C^\infty(U,E) \to C(U \times E, E)$ is continuous. We can thus write the adjoint map as a composition of continuous mappings (see Lemma B.10, which uses the that U is locally compact, i.e. E finite dimensional) $p(f,g) = \mathrm{ev}(d(f),u, \mathrm{ev}(g,u))$, whence the Lie bracket is continuous. □

D.13 Corollary *Let M be a finite-dimensional manifold. Then $(\mathcal{V}(M),[\cdot,\cdot])$ is a locally convex Lie algebra.*

Proof That the vector fields form a Lie algebra is checked in Exercise 3.2.3. Recall from D.4 that the vector fields were topologised as a subspace of a product of spaces of the form $C^\infty(U,E)$, where $U \subseteq M$. By construction of the Lie bracket of two vector fields, the bracket is given by a local formula on chart domains U. Hence it suffices to establish continuity of the local formula on the spaces $C^\infty(U,E)$. This was exactly the content of Lemma D.12. □

D.14 Remark In general, the Lie algebra of vector fields $\mathcal{V}(M)$ will not be a locally convex Lie algebra if M is an infinite-dimensional manifold. Indeed, it can be shown that Lemma D.12 becomes false beyond the realm of Banach spaces. To see this, let $U \subseteq E$ be an open subset of a non-normable space. We consider the subalgebra

$$\begin{aligned}\mathcal{A} = \{X_{A,b} \in C^\infty(U,E) \mid &\text{ for all } v \in E,\ X_{A,b}(v)\\ &= Av + b, \text{ for } A \in L(E,E), b \in E\}\end{aligned}$$

of affine vector fields. By construction we can identify $\mathcal{A} \cong L(E,E) \times E$. Here the subspace topology induced by the compact-open C^∞-topology of $C^\infty(U,E)$ on $\mathcal{A}$ is the product topology, where E carries its natural locally convex topology and the space of continuous linear mappings $L(E,E)$ is endowed with the compact-open topology (i.e. the topology induced by the embedding $L(E,E) \subseteq C_{\mathrm{c.o.}}(E,E)$). Indeed the latter fact is irrelevant for us; we are only interested in the fact that this topology turns $L(E,E)$ into a topological vector space. Now, the Lie bracket of $C^\infty(U,E)$ induces the Lie bracket

$$[X_{A,b}, X_{C,d}](v) = (A \circ C - C \circ A)(v) + (A(d) - C(b))$$

on the affine vector fields (these facts are left as Exercise D.1.3). To see that this Lie bracket is, in general, not continuous, it suffices to note that the evaluation map $L(E,E) \times E \to E$, $(A,v) \mapsto A(v)$ is discontinuous. For this we pick $0 \neq v \in E$ and consider the mapping

$$j\colon E' = L(E,\mathbb{R}) \to L(E,E), \quad \lambda \mapsto (x \mapsto \lambda(x) \cdot v). \tag{D.5}$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

If we endow the dual space E' with the compact-open topology (again the subspace topology of $E' \subseteq C_{\text{c.o.}}(E,\mathbb{R})$) then E' becomes a topological vector space and j continuous. However, Proposition A.19 shows that the evaluation map $E' \times E \to \mathbb{R}$ is discontinuous for every topological vector space E which is not normable. As j and scalar multiplication in E are continuous, this implies that the evaluation of $L(E,E)$ must be discontinuous if E is not normable. We deduce that the Lie bracket on $C^\infty(U,E)$ must be discontinuous if E is not normable.[1]

Exercises

D.1.1 Show that the construction of the topology for $\mathcal{V}(M)$ in D.4 is just a special case of C.7.

D.1.2 Let A be an associative algebra. Show that the set of derivations $\operatorname{der}(A)$ (see Definition D.10) forms a Lie subalgebra of $(L(A,A),[\cdot,\cdot])$, where the bracket is given by the commutator bracket $[f,g] = f \circ g - g \circ f$ of linear maps.

D.1.3 We provide the missing details in Remark D.14. To this end let $U \subseteq\!\!\!\circ\, E$ in a locally convex space and endow $C^\infty(U,E)$ with the compact-open topology (i.e. the topology induced by the embedding $L(E,E) \subseteq C_{\text{c.o.}}(E,E)$). We consider the affine vector fields

$$\mathcal{A} = \{X_{A,b} \in C^\infty(U,E) \mid \text{for all } v \in E,\ X_{A,b}(v) = Av + b, \\ \text{for } A \in L(E,E),\ b \in E\}$$

and identify $\mathcal{A} = L(E,E) \times E$ (where $L(E,E)$ denote continuous linear maps). Show that:

(a) The subspace topology on $\mathcal{A}$ is the product topology of the compact-open topology on $L(E,E)$ and the locally convex topology of E.

(b) The Lie bracket on $C^\infty(U,E)$ induces the Lie bracket

$$[X_{A,b}, X_{C,d}] = (A \circ C - C \circ A) + (A(d) - C(b)) \text{ on } \mathcal{A}.$$

(c) If we endow the dual space E' with the compact-open topology (i.e. the subspace topology of $E' \subseteq C_{\text{c.o.}}(E,\mathbb{R})$), then E' is a topological vector space and the map j from (D.5) becomes continuous.

[1] Even stronger, one can show that the evaluation must be discontinuous on $L(E, E)$ with the compact-open topology for all infinite-dimensional spaces E; see Neeb (2006, Remark I.5.3) for an exposition.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Appendix E
Differential Forms on Infinite-Dimensional Manifolds

E.1 Introduction

In this appendix we give a short introduction to differential forms on infinite-dimensional manifolds. For more information on differential forms on infinite-dimensional manifolds and their application, we refer the interested reader to Beggs (1987), as well as Glöckner and Neeb (forthcoming). The main difference between the finite-dimensional (or Banach) and our setting is that it is, in general, impossible to interpret differential forms as (smooth) sections into certain bundles of linear forms. The reason for this is again that the topology on spaces of linear forms breaks down beyond the Banach setting (see Proposition A.19). Even worse, the many equivalent ways to define differential forms in finite dimensions become inequivalent in the infinite-dimensional setting (see Kriegl and Michor, 1997, Section 33 for a thorough discussion of this phenomenon). Most notably, there is no useful way to describe differential forms as a sum of differential forms coming from a local coordinate system.

We begin with the definition of a differential form. This definition is geared towards avoiding any reference to topologies on spaces of linear mappings. This again is a continuity problem and in the inequivalent convenient setting of global analysis, differential forms can be described as sections in suitable bundles; see Kriegl and Michor (1997, 33.22 Remark). Furthermore, we need to avoid arguments involving the existence of (smooth) bump functions (which in general do not exist; see Appendix A.4).

E.1 Definition Let M be a manifold and E be a locally convex space and $p \in \mathbb{N}_0$. An *E-valued p-form* ω on M is a function ω which associates to each $x \in M$ a p-linear alternating map $\omega_x : (T_x M)^p \to E$ such that for each chart (U, φ) of M, the map

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$\omega_\varphi \colon V_\varphi \times F^p_\varphi \to E, \quad \omega_\varphi(x, v_1, \dots, v_p) \coloneqq \omega_{\varphi^{-1}(x)}(T_x\varphi^{-1}(v_1), \dots, T_x\varphi^{-1}(v_p)) \tag{E.1}$$

is smooth. We write $\Omega^p(M,E)$ for the *space of smooth E-valued p-forms* on M. Note that $\Omega^0(M,E) = C^\infty(M,E)$.

E.2 Example For $p = 0$, we have already seen that smooth functions are differential forms. If $f \in C^\infty(M,E)$, then the derivative $df \colon TM \to E$, $v \mapsto \mathrm{pr}_2 \circ Tf(v)$ is a smooth E-valued 1-form.

Constructing the derivative of a smooth function can be generalised to a differential on the space of p-forms, the so-called *exterior differential*

$$\mathrm{d} \colon \Omega^p(M,E) \to \Omega^{p+1}(M,E)$$

which we discuss now. On an infinite-dimensional manifold there is no generalisation of local coordinates in a vector basis. Hence the finite-dimensional approach defining the exterior differential in a local coordinate frame is not available. Recall some standard notation useful in the present context: In (D.1) we defined a derivative $X.f$ of a smooth function f in the direction of X. If $\omega \in \Omega^p(M,E)$ and $U \subseteq M$, we define for vector fields $X_1, \dots X_p \in \mathcal{V}(U)$ a smooth map

$$\omega(X_1, \dots, X_p) \colon U \to E, \quad m \mapsto \omega_m(X_1(m), \dots X_p(m)).$$

Finally, we write $\omega(X_0, \dots, \hat{X}_i, \dots, X_p)$ to indicate that the ith component is to be omitted from the formula.

E.3 Proposition *For $\omega \in \Omega^p(M,E)$ there exists a smooth $p+1$-form $\mathrm{d}\omega \in \Omega^{p+1}(M,E)$ which for any $U \subseteq M$ and vector fields $X_1, \dots, X_p \in \mathcal{V}(U)$, satisfies*

$$\begin{aligned} \mathrm{d}\omega(X_0, X_1, \dots, X_p)(m) &= \sum_{i=0}^{p} (-1)^i (X_i.\omega(X_0, \dots, \hat{X}_i, \dots, X_p))(m) \\ &+ \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_0, \dots, \hat{X}_i, \dots \hat{X}_j, \dots, X_p)(m). \end{aligned} \tag{E.2}$$

Proof Consider $m \in M$ and $v_1, \dots, v_p \in T_m M$. To define $(\mathrm{d}\omega)_m(v_1, \dots, v_p)$ we pick an open neighbourhood U of m together with vector fields $X_i \in \mathcal{V}(U)$ such that $X_i(m) = v_i, i = 0, 1, 2, \dots, p$. Note that such vector fields always exist as we can take the constant vector fields in a chart neighbourhood (in particular, the definition does not require us to globalise these fields, which would require bump functions which may not exist). Then

$$(\mathrm{d}\omega)_m(v_0, v_1, \dots, v_p) \coloneqq \mathrm{d}\omega(X_0, X_1, \dots, X_p)(m), \tag{E.3}$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

where the right-hand side has been defined via (E.2) for our choice of vector fields.

Step 1: $(\mathrm{d}\omega)_m(v_1,\dots,v_p)$ does not depend on the choice of vector fields in (E.3). We have to show that the expression (E.2) becomes 0 if $X_k(m) = 0$ for at least one k. Assuming that X_k vanishes in m, we may without loss of generality assume that we are working in local coordinates. We will suppress the chart identification in the formulae and also identify each vector field X_i with its principal part on some $U \subseteq F$ (where F is a locally convex space). Exploit now that ω is alternating and linear to see that the contributions in (E.2) which do not directly vanish are

$$\sum_{i\neq k}^{p}(-1)^i(X_i.\omega(X_0,\dots,\hat{X}_i,\dots,X_p))(m) \tag{E.4}$$

$$+\sum_{i<k}(-1)^{i+k}\omega([X_i,X_k],X_0,\dots,\hat{X}_i,\dots,\hat{X}_k,\dots,X_p)(m) \tag{E.5}$$

$$+\sum_{k<i}(-1)^{i+k}\omega([X_k,X_i],X_0,\dots,\hat{X}_k,\dots,\hat{X}_i,\dots,X_p)(m). \tag{E.6}$$

Apply the definition of the differential form ω on the open subset of a locally convex space (E.1). In this presentation ω is a function of $p+1$-variables and p-linear in the last p-variables. Hence we can compute the derivative for a summand in (E.4) explicitly as

$$\begin{aligned}
&X_i.\omega(X_0,\dots,\hat{X}_i,\dots,X_p)(m)\\
&\quad= d_1\omega(m,X_1(m),\dots\hat{X}_i(m),\dots,X_p(m);X_i(m))\\
&\qquad+\sum_{j<i}\omega_m(X_0(m),\dots,dX_j(m;X_i(m)),\dots,\hat{X}_i(m),\dots,X_p(m))\\
&\qquad+\sum_{i<j}\omega_m(X_0(m),\dots,\hat{X}_i(m),dX_j(m;X_i(m)),\dots,X_p(m)).
\end{aligned}$$

As X_k vanishes, we see that for every $i > k$ only

$$\omega_m(X_0(m),\dots,dX_k(m;X_i(m)),\dots,\hat{X}_i(m),\dots,X_p(m))$$

survives and as $X_k(m) = 0$, we have

$$dX_k(m,X_i(m)) = dX_k(m;X_i(m)) - dX_i(m;X_k(m)) = [X_i,X_k](m).$$

Using the fact that ω is alternating, we have

$$\begin{aligned}
&(-1)^k\omega([X_k,X_i],X_0,\dots,\hat{X}_k,\dots,\hat{X}_i,\dots,X_p)(m)\\
&\quad=-\omega_m(X_0(m),\dots,dX_k(m;X_i(m)),\dots,\hat{X}_i(m),\dots,X_p(m)),
\end{aligned}$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

hence the corresponding terms in (E.4) and (E.5) cancel. Similar arguments show that this also happens for the parts in (E.4) and (E.6) if $i < k$. We conclude that (E.2) vanishes at a point if one of the vector fields vanishes at the point. This shows, in particular, that (E.3) is independent of the choices of vector fields.

Step 2: $\mathrm{d}\omega$ is a smooth $p+1$-form. For smoothness we work again locally in a chart as in Step 1 and pick all vector fields X_i to be constant. As the Lie bracket of constant vector fields vanishes, (E.3) reduces to

$$(\mathrm{d}\omega)_m(v_0,\dots,v_p) = \sum_{i=0}^{p}(-1)^i d_1\omega(m,v_0,\dots,\hat{v}_i,\dots,v_p;v_i). \tag{E.7}$$

Now by definition ω induces a smooth function in all charts, whence we see that $\mathrm{d}\omega$ is also smooth in $(m,v_0,v_1,\dots,v_o)$. To see that $(\mathrm{d}\omega)_m$ is alternating, observe that the summands in (E.7) are alternating. Assume now that $v_i = v_j$ for some $i < j$. It is easy to see that (E.7) vanishes (we leave this as Exercise E.1.1). □

We thus obtain for every $p \geq 0$ an exterior differential on the space of p-forms. The usual proof (see Kriegl and Michor, 1997, Theorem 33.18, or Lang, 1999, V. Proposition 3.3) then shows that $\mathrm{d}^2 = \mathrm{d}\circ\mathrm{d} = 0$ in every degree. Hence as in the finite-dimensional (or the Banach) setting, the exterior differential gives rise to a cochain complex of differential forms

$$C^\infty(M,E) = \Omega^0(M,E) \xrightarrow{\mathrm{d}} \Omega^1(M,E) \xrightarrow{\mathrm{d}} \Omega^3(M,E) \xrightarrow{\mathrm{d}} \cdots .$$

Starting from this complex, one can define and study de Rham cohomology on the (infinite-dimensional) manifold M. We will not pursue this route here and refer instead to Beggs (1987) or Kriegl and Michor (1997, Chapter 34) for more information.

Differential forms of higher order are typically constructed using the wedge product. The definition is as in the finite-dimensional setting (note, however, that there are several conventions as to the coefficients; we chose to follow Lang, 1999).

E.4 Definition Let E_i, $i = 1,2,3$ be locally convex spaces and $\beta\colon E_1 \times E_2 \to E_3$ be a continuous bilinear map. Fix $p,q \in \mathbb{N}_0$ and denote by S_{p+q} the symmetric group of all permutations of $\{1,2,\dots p+q\}$. For $\omega \in \Omega^p(M,E_1)$ and $\eta \in \Omega^q(M,E_2)$, define the *wedge product* $\omega \wedge \eta \in \Omega^{p+q}(M,E_3)$ via $(\omega\wedge\eta)_x := \omega_x \wedge \eta_x$ for $x \in M$, where

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$(\omega_x \wedge \eta_x)(v_1,\dots,v_{p+q})$$
$$:= \frac{1}{p!q!}\sum_{\sigma\in S_{p+q}} \operatorname{sgn}(\sigma)\beta(\omega_x(v_{\sigma(1)},\dots,v_{\sigma(p)}),\eta_x(v_{\sigma(p+1)},\dots,v_{\sigma(p+q)})).$$

Then

$$\wedge\colon \Omega^p(M,E_1)\times\Omega^q(M,E_2)\to\Omega^{p+q}(M,E_3),\quad (\omega,\eta)\mapsto\omega\wedge\eta$$

is a bilinear map.

E.5 Example If $s\colon \mathbb{R}\times E\to E$ is the scalar multiplication and $f\in C^\infty(M,\mathbb{R})$, then the wedge product of f and $\omega\in\Omega^p(M,E)$ is given by $(f\wedge\omega)_x = f(x)\omega_x$. This is usually abbreviated by $f\omega := f\wedge\omega$ and it is easy to see that $\Omega^p(M,E)$ becomes a $C^\infty(M,\mathbb{R})$-module.

E.6 Example Let $(E,[\cdot,\cdot])$ be a locally convex Lie algebra. Then $[\cdot,\cdot]\colon E\times E\to E$ is bilinear and we can construct the wedge product $\wedge$ with respect to the Lie bracket. For this special situation we define for $\omega\in\Omega^p(M,E)$ and $\eta\in\Omega^q(M,E)$ the bracket

$$[\omega,\eta]_\wedge := \omega\wedge\eta.$$

We will now define several standard operations on differential forms such as the pullback of p-forms by smooth mappings.

E.7 Definition Let $\varphi\colon M\to N$ be a smooth map between manifolds. Then we define for $\omega\in\Omega^p(N,E)$ a p-form $\varphi^\omega\in\Omega^p(M,E)$, the *pullback of ω by φ* via

$$(\varphi^*\omega)_x(v_1,\dots v_p) := \omega_{\varphi(x)}(T_x\varphi(v_1),\dots T_x\varphi(v_p)).$$

Due to the chain rule we immediately have the following rules for the computation of pullbacks.

E.8 The following rules hold for smooth maps and p-forms

$$\mathrm{id}_M^*\,\omega=\omega,\quad \varphi_1^*(\varphi_2^*\omega)=(\varphi_2\circ\varphi_1)\omega^*,\quad \varphi^*(\omega\wedge\eta)=\varphi^*\omega\wedge\varphi^*\eta.$$

If $p=0$, that is, $\omega=f\in C^\infty(M,E)$, then $\varphi^*f=f\circ\varphi$ and we recover the pullback discussed in the context of manifolds of mappings.

Finally, we define the Lie derivative of a differential form by a vector field. Before we begin, note that the definition of the Lie derivative has to diverge from the usual definition on finite-dimensional or Banach manifolds. This is due to the fact that the common description of the Lie derivative (see e.g. Lang, 1999, V. §2) uses the differential of a flow of a vector field. However, as flows of vector fields are the solutions to certain ordinary differential equations, it is unclear whether the flow of a vector field would exist on the more general

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

manifolds we consider (see Appendix A.6 for a discussion of this problem). Nevertheless, Lang (1999, V.5 Proposition 5.1) shows that for Banach manifolds the following definition coincides with the classical one involving flows.

E.9 Definition Let M be a manifold and E a locally convex space. For $X \in \mathcal{V}(M)$ and $\omega \in \Omega^p(M,E), p \in \mathbb{N}_0$ we define the *Lie derivative* $\mathcal{L}_Y\omega \in \Omega^p(M,E)$ as follows:

$$\begin{aligned}(\mathcal{L}_Y\omega)_m(v_1,\dots,v_p)&\\ &= Y.\omega(X_1,\dots,X_p)(m) - \sum_{j=1}^{p}\omega(X_1,\dots,[Y,X_j],\dots X_p)(m)\\ &= Y.\omega(X_1,\dots,X_p)(m) + \sum_{j=1}^{p}(-1)^j\omega([Y,X_j],X_1,\dots,\hat{X}_j,\dots X_p)(m),\end{aligned}$$

where the X_i are smooth vector fields defined in a neighbourhood of m such that $X_i(m) = v_i$. That the Lie derivative is well defined will be checked in Exercise E.1.4.

Note that for $\omega \in \Omega^0(M,E) = C^\infty(M,E)$ the formula of the Lie derivative reduces to $\mathcal{L}_Y\omega = d\omega \circ Y$. This was precisely the formula for the Lie derivative described in Definition D.10 for functions.

Exercises

E.1.1 Check that the exterior differential $\mathrm{d}\omega$ of a p-form is an alternating $p+1$-form.

E.1.2 Check the details in Definition E.4. Show that

(a) the wedge product of a p-form and a q-form is indeed a $p+q$-form;

(b) the wedge product defines a bilinear map between spaces of differential forms;

(c) $\Omega^p(M,E)$ is a $C^\infty(M,\mathbb{R})$-module (see Example E.5);

(d) for $p = q = 1$ we have $\omega_x \wedge \eta_x(v_1,v_2) = \beta(\omega_x(v_1),\eta_x(v_2)) - \beta(\omega_x(v_2),\eta_x(v_2))$.

E.1.3 Prove that for $\omega \in \Omega^p(M,E)$, $\eta \in \Omega^q(M,F)$ and any wedge product, the following formula holds:

$$\mathrm{d}(\omega \wedge \eta) = (\mathrm{d}\omega) \wedge \eta + (-1)^p\omega \wedge (\mathrm{d}\eta).$$

Furthermore, show that for $f\colon N \to M$ smooth, we have

$$f^*\mathrm{d}\omega = \mathrm{d}f^*\omega.$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Hint: The assertions are local, whence they can be solved using the local formula for the exterior differential.

E.1.4 In this exercise we let $X \in \mathcal{V}(M)$ and $\omega \in \Omega^p(M,E), p \in \mathbb{N}_0$ for M a manifold and E a locally convex space. Show that the definition of the Lie derivative $\mathcal{L}_Y\omega$ does not depend on the choice of vector fields X_i in Definition E.9. Conclude that $\mathcal{L}_Y\omega$ is a smooth p-form.
Hint: It suffices to show that $\mathcal{L}_Y\omega$ vanishes if $X_i(m) = 0$, and this can be checked locally.

E.2 The Maurer–Cartan Form on a Lie Group

For Lie groups there are two important differential forms induced by the Lie group structure.

E.10 Example Let G be a Lie group with Lie algebra $\mathbf{L}(G)$. Then we define the *right Maurer–Cartan form* $\kappa^r \in \Omega^1(G,\mathbf{L}(G))$ via

$$(\kappa^r)_g : T_gG \to \mathbf{L}(G), \quad v \mapsto T_g\rho_{g^{-1}}(v),$$

where $\rho_{g^{-1}}(h) = hg^{-1}$.

Now if $f \in C^\infty(M,G)$, we can define its right logarithmic derivative via

$$\delta^r f \colon M \to \mathbf{L}(G), \quad \delta^r f := f^*\kappa^r.$$

Similarly, one can define the *left Maurer–Cartan form* $\kappa^\ell \in \Omega^1(G,\mathbf{L}(G))$ and a left logarithmic derivative by replacing right multiplication with left multiplication in the definition of κ^r. This generalises the construction of the logarithmic derivatives for curves from 3.31. One can show (see Exercise E.2.1) that the left logarithmic derivative of any function satisfies the right Maurer–Cartan equation

$$\mathrm{d}\delta^\ell f + \frac{1}{2}[\delta^\ell f, \delta^\ell f]_\wedge = 0, \tag{E.8}$$

where $[\delta^\ell f, \delta^\ell f]_\wedge = \delta^\ell f \wedge \delta^\ell f$ for the wedge product induced by the Lie bracket.

E.11 Definition Let G be a Lie group with Lie algebra $\mathbf{L}(G)$ and $\omega \in \Omega^1(M,\mathbf{L}(G))$ for some smooth manifold M. Then ω is called

(a) *integrable* if there exists $f \in C^\infty(M,G)$ with $\omega = \delta^\ell f$;
(b) *locally integrable* if for every $m \in M$ there is $m \in U \subseteq M$ such that $\omega|_U$ is integrable.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Equivalently we could have defined (local) integrability using the right logarithmic derivative. With this definition it is possible to formulate the fundamental theorem for Lie group-valued functions with values in regular Lie groups; see §3.3.

E.12 Proposition (Fundamental theorem for Lie group-valued functions) *Let G be a regular Lie group with Lie algebra $\mathbf{L}(G)$ and $\omega \in \Omega^1(M,\mathbf{L}(G))$. If ω satisfies the right Maurer–Cartan equation* (E.8)*, then ω is locally integrable. If, in addition, M is simply connected, then ω is integrable.*

The proof of Proposition E.12 needs concepts (e.g. connections on principal bundles) which we will not introduce here. Instead we refer the interested reader either to the classical proofs for the finite-dimensional setting, for example, Sharpe (1997, 3.§6–7), or to the infinite-dimensional sources Neeb (2006, Theorem III.2.1) as well as Kriegl and Michor (1997, Theorem 40.2).

E.13 Lemma *Let M be a connected manifold and $\varphi,\psi \in C^\infty(M,G)$, where G is a Lie group. Then $\delta^\ell\varphi = \delta^\ell\psi$ is equivalent to the existence of $g \in G$ with $\varphi = g \cdot \psi$.*

Proof If there exists $g \in G$ with $\varphi = g \cdot \psi$, then a straightforward calculation shows that $\delta^\ell\varphi = \delta^\ell\psi$ holds. Assume conversely that $\delta^\ell\varphi = \delta^\ell\psi$. Then define the map $\Psi := \varphi \cdot \psi^{-1}$ (where the product and inverse are taken pointwise in G). Exercise E.2.3 yields

$$\delta(\gamma) = T_e\lambda_\psi T_e\rho_{\psi^{-1}}(\delta^\ell\varphi - \delta^\ell\psi) = 0,$$

or in other words, the map γ is locally constant by Corollary 1.19. As M is connected, we conclude that $g := \gamma(m) \in G$ (for any $m \in M$) satisfies $g \circ \psi = \varphi$. □

E.14 Proposition (Lie II for regular Lie groups) *Let G,H be Lie groups with Lie algebras $\mathbf{L}(G)$ and $\mathbf{L}(H)$, respectively. Let $f\colon \mathbf{L}(G) \to \mathbf{L}(H)$ be a morphism of locally convex Lie algebras. If H is a regular Lie group and G is connected and simply connected, then there exists a unique morphism of Lie groups $\varphi\colon G \to H$ with $\mathbf{L}(\varphi) = f$.*

Proof Since f is continuous we can consider the smooth 1-form $\alpha := f \circ \kappa^\ell \in \Omega^1(G,\mathbf{L}(H))$, where κ^ℓ is the left Maurer–Cartan form on G. Moreover, since the Maurer–Cartan form is left invariant, so is α. We consider the Maurer–Cartan equation on H (to mark this we label the bracket operation $[\cdot,\cdot]_\wedge^H$ by H). As the wedge is induced by the Lie bracket on $\mathbf{L}(H)$, we can exploit that f is a Lie algebra morphism and compute with Exercise E.1.3 as follows:

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$\mathrm{d}\alpha+\frac{1}{2}[\alpha,\alpha]^H_\wedge=\mathrm{d}(f\circ\kappa^\ell)+\frac{1}{2}[f\circ\kappa^\ell,f\circ\kappa^\ell]^H_\wedge=f\circ\left(\mathrm{d}\kappa^\ell+\frac{1}{2}[\kappa^\ell,\kappa^\ell]_\wedge\right)=0,$$

where the unlabelled bracket is the one induced by the Lie bracket of $\mathbf{L}(G)$ and we exploited that the Maurer–Cartan form satisfies the Maurer–Cartan equation on G by Exercise E.2.1. Now the fundamental theorem, Proposition E.12, implies that there is a mapping $\varphi\colon G\to H$ with $\delta^\ell\varphi=\alpha$. Fixing $\varphi(e_G)=e_H$, this mapping is unique by Lemma E.13. Consider now $g\in G$ and $\varphi\circ\lambda_g$. Then we pick $v\in T_kG$ and evaluate the differential form

$$\delta^\ell(\varphi\circ\lambda_g)(v)=T\lambda_{\varphi(gk)^{-1}}T_{gk}\varphi T\lambda_g(v)=\delta^\ell\varphi(T\lambda_g(v))=\lambda_g^*\alpha=\alpha.$$

Now applying Lemma E.13 again, the maps $\varphi\circ\lambda_g$ and φ differ only by left translation with an element which we compute as $\varphi\circ\lambda(e_G)=\varphi(g)$. In other words, $\varphi(gk)=\varphi(g)\varphi(k)$ for all $k\in G$. Since $g\in G$ was arbitrary, we see that φ is indeed a morphism of Lie groups. □

Exercises

E.2.1 Let G be a Lie group with Lie algebra $\mathbf{L}(G)$ and left Maurer–Cartan form κ^ℓ. Show that κ^ℓ:

(a) is an $\mathbf{L}(G)$-valued differential form on G which is left invariant in the sense that $\lambda_g^*\kappa^\ell=\kappa^\ell$ for each $g\in G$ (where $\lambda_g(h)\coloneqq gh$;

(b) satisfies the right Maurer–Cartan equation

$$\mathrm{d}\kappa^\ell+\frac{1}{2}[\kappa^\ell,\kappa^\ell]_\wedge=0.$$

Hint: Compute the exterior derivative locally using (E.3). It suffices to prove the formula using left invariant vector fields (why?).

Remark: The 'right Maurer–Cartan equation' is related to the right principal action of G on itself by multiplication. There is also a corresponding left Maurer–Cartan equation for the right Maurer–Cartan form, where the bracket in the equation gets a negative sign.

(c) Deduce that for a smooth function f, $\delta^\ell f$ also satisfies the right Maurer–Cartan equation.

E.2.2 Let $\varphi\colon G\to H$ be a morphism of Lie groups and κ^ℓ the (left) Maurer–Cartan form on G. Show that:

(a) $\delta^\ell\varphi=\mathbf{L}(\varphi)\circ\kappa^\ell$;

(b) if $\psi\colon G\to H$ is another Lie group morphism with $\delta^\ell\varphi=\delta^\ell\psi$, then $\varphi=\psi$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

E.2.3 Let $f, g \in C^\infty(M, G)$ be smooth maps to a Lie group. Establish the following quotient rule for the left logarithmic derivative:

$$\begin{aligned}\delta^\ell(f \cdot g^{-1})(m) &= T_{\mathbf{1}_G}\lambda_{g(m)} T_{\mathbf{1}_G}\rho_{g^{-1}(m)}(\delta^\ell f(m) - \delta^\ell g(m)) \\ &= \mathrm{Ad}_{g(m)}(\delta^\ell f(m) - \delta^\ell g(m)),\end{aligned}$$

where $\mathbf{1}_G \in G$ is the identity element, products and inverses are taken pointwise and λ (resp. ρ) denotes left (resp. right) multiplication in the Lie group.
Hint: Apply Lemma 3.12.

E.3 Supplement: Volume Form and Classical Differential Operators

In this short supplement we will record some well-known facts on differential forms on finite-dimensional (compact) manifolds. Many of these notions are needed in Chapter 7 and we recall them for the reader's convenience. Thus detailed proofs will, in general, be omitted in this section. However, all of these results are readily available in the standard finite-dimensional literature (which we will reference).

Conventions We fix (M, g) a compact (thus finite-dimensional) and connected manifold with Riemannian metric g. Furthermore, we denote by $d = \dim M$ the dimension of the model space of M.

Let us first recall that there is another canonical way to define differential forms.

E.15 (Differential forms as sections; see Klingenberg, 1995, 1.4; Abraham et al., 1988, Section 6) Starting with the tangent bundle TM we can construct the bundle of alternating k-forms $\mathcal{A}^k(M) \to M$ for $k \geq 0$. The fibre of $\mathcal{A}^k(M)$ over $x \in M$ is given by the space of alternating k-linear mappings $(T_xM)^k \to \mathbb{R}$ which we denote by $L^k_a(T_xM, \mathbb{R})$ and topologise as a subspace of the k-linear mappings (which carry the usual norm topology induced by the operator norm for k-linear maps; see Lang, 1999, I. §2). Further, every chart (U, φ) of M induces a vector bundle trivialisation of $\mathcal{A}^k(M)$ over U via

$$\kappa_\varphi(x, \omega) := (\varphi(x), \omega \circ (T_x\varphi^{-1} \times T_x\varphi^{-1} \times \cdots \times T_x\varphi^{-1})).$$

Comparing the construction with (E.1), it becomes clear that for the finite-dimensional manifold M (indeed for any Banach manifold) differential k-forms are just smooth sections of $\mathcal{A}^k(M)$. In other words, we obtain

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$$\Omega^k(M) = \Gamma(\mathcal{A}^k(M)), \quad \text{for all } k \in \mathbb{N}_0.$$

Moreover, this allows us to topologise the space of differential k-forms as a locally convex vector space via C.7. Namely if we pick a family of open sets $(U_i)_{i\in I}$ which covers M such that on each U_i there is a bundle trivialisation $(\kappa_i)^{-1}\colon \mathcal{A}^k(M)|_{U_i} \to U_i \times \mathbb{R}^d$, $x \mapsto ((\kappa_i^1)^{-1}(x), (\kappa_i^2)^{-1}(x))$, then the mapping

$$\Omega^k(M) \to \prod_{i\in I} C^\infty(U_i, L_a^k(\mathbb{R}^d, \mathbb{R})), \quad \omega \mapsto \omega \circ (\kappa_i^1, \kappa_i^2, \ldots, \kappa_i^2)$$

is an embedding of $\Omega^k(M)$ as a closed locally convex subspace of the product on the right-hand side.

Recall that the dimension of the spaces $L_a^k(T_xM, \mathbb{R})$ depends on the dimension $d = \dim T_xM$. In particular, $L_a^d(T_xM, \mathbb{R})$ is one-dimensional and a differential form $\mu \in \Omega^d(M)$ which vanishes nowhere is called a *volume form*. On Riemannian manifolds there is a convenient way to construct a volume form associated to the Riemannian metric. Recall that a manifold M is orientable, if it admits an atlas $(U_i, \varphi_i)_{i\in I}$ such that the Jacobians of all change of charts are positive. One can prove (Gallot et al., 2004, Theorem 1.127) that a Riemannian manifold is orientable if and only if it admits a volume form μ (induced by the Riemannian metric). Volume forms are the tool of choice to define integration on manifolds; see Lang (1999, Part III) or Abraham et al. (1988, Section 7). In particular, we can define the L^2-metric on $\mathcal{V}(M)$ in the presence of a volume form symbolically without defining the integral as follows:

$$g_{L^2}(X,Y) = \int_M g(X,Y)\mathrm{d}\mu, \qquad X,Y \in \mathcal{V}(M). \tag{E.9}$$

In Chapter 5 we often considered only integration on $\mathbb{S}^1$ since a global parametrisation allowed us to hide the dependence on a volume form and (E.9) reduces for $M = \mathbb{S}^1$ to (5.1).

Classical Differential Operators on a Riemannian Manifold

We will now assume that there is a volume form μ associated to the Riemannian metric on M. Let us then recall the following classical differential operators on M.

E.16 Definition (Abraham et al., 1988, Sections 6.5 and 7.5) For a compact Riemannian manifold (orientable in case we need a volume form μ), we will consider the following differential operators.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

- For a vector field X, there is a unique smooth function $\operatorname{div} X\colon M \to \mathbb{R}$, the *divergence* of X such that

$$\mathcal{L}_X \mu = (\operatorname{div} X)\mu.$$

- If $f \in C^\infty(M,\mathbb{R})$ we exploit that the Riemannian metric induces an isomorphism $TM \cong T^*M$ (see Proposition 4.5). Thus the following formula uniquely determines the *gradient of f with respect to g*:

$$g_m(\operatorname{grad} f(m), v_m) = df(v_m) \quad \text{for all } m \in M, v_m \in T_m M.$$

- The *(Hodge)Laplacian* $\Delta = \mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d}$ is associated to the metric (see Lang, 1999, p. 423). Here d is the exterior differential from Proposition E.3 and d^* is the codifferential defined via the Hodge star (this is a finite-dimensional construction which depends on the Riemannian metric g; we refer to Lee, 2013, p. 464 for more information).

Having defined the necessary differential operators, we recall two decompositions which are, for example, relevant in geometric hydrodynamics.

E.17 Proposition (Helmholtz decomposition, see Modin (2019), Lemma 1.2) *Let (M,g) be a compact oriented Riemannian manifold with volume form μ and $X \in \mathcal{V}(M)$. Then there exist $V \in \mathcal{V}(M)$ and $f \in C^\infty(M,\mathbb{R})$ such that*

$$X = V + \operatorname{grad} f \qquad \textit{and} \quad \operatorname{div} V = 0.$$

Moreover, V and $\operatorname{grad} f$ *are orthogonal with respect to the L^2-metric* (E.9)*, that is,*

$$g_{L^2}(V, \operatorname{grad} f) = \int_M g(V, \operatorname{grad} f)\mathrm{d}\mu = 0.$$

Note that since the differential d and the codifferential d^* make sense for arbitrary k-forms, we can also extend the Hodge Laplacian to k-forms. This induces the Hodge decomposition of k-forms; see, for example, Taylor (2011, Proposition 8.2). We will not recall it here, but would like to mention that it is an important ingredient to establish the Lie group structure of the groups of volume-preserving diffeomorphisms and the symplectomorphism group; Example 3.10. As we now have the necessary notation in place, let us very briefly sketch the idea of the proof.

E.18 (Submanifold structure of volume-preserving diffeomorphisms (sketch)) Define the map $\Psi_\mu\colon \operatorname{Diff}(M) \to \Omega^d(M)$, $\phi \mapsto \phi^*\mu$. Since M is compact, we endow $\Omega^d(M)$ via E.15 with a locally convex vector space structure. With some work one can show that Ψ_μ is a smooth map with derivative $T_\phi\Psi_\mu(V_\phi) = \phi^*(\mathcal{L}_{V_\phi}\omega)$ (this follows somewhat similarly to the proof that the pullback with

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

smooth functions is smooth). Then one needs to prove that Ψ_μ is a submersion onto the cohomology class $[\mu] = \mu + \mathrm{d}\Omega^{d-1}(M) \subseteq \Omega^d(M)$ of μ. The proof uses the Hodge decomposition of d-forms to construct a splitting of the kernel of Ψ_μ. Further, one needs to work in a Sobolev completion of $\mathrm{Diff}(M)$, whence this is beyond the techniques we are developing in this book. We refer the interested reader to Smolentsev (2007), along with Ebin and Marsden (1970). Then the volume-preserving diffeomorphism group is simply the preimage $\Psi_\mu^{-1}(\mu)$ of the singleton μ. In particular, $\mathrm{Diff}_\mu(M)$ is a submanifold and thus a Lie subgroup of $\mathrm{Diff}(M)$.

Note that it is apparent from the derivative of Ψ_μ and Exercise 1.7.4 that the Lie algebra of $\mathrm{Diff}_\mu(K)$ is $\mathcal{V}_\mu(M) = \{X \in \mathcal{V}(M) \mid \operatorname{div} X = 0\}$, the Lie algebra of divergence-free vector fields.

Exercises

E.3.1 Show that the structure described in E.15 yields a vector bundle $\mathcal{A}^k(M) \to M$.

E.3.2 Prove that the characterisation of differential forms via the bundle in E.15 coincides with the one from Definition E.1.

E.3.3 Consider $\mathbb{R}^d$ as a Riemannian manifold with the standard Euclidean metric. Convince yourself that div, grad and Δ are 'the usual' differential operators from vector calculus in this case.

E.3.4 Work out the details for E.18 (note that this requires the Hodge decomposition theorem, Abraham et al., 1988, Theorem 7.5.3). Show that:

(a) Ψ_μ is smooth with surjective derivative;

(b) the kernel of $T_\eta \Psi_\mu$ is a split subspace of $T_\eta \mathrm{Diff}(M)$.

Remark: If $\mathrm{Diff}(M)$ were a Banach manifold the above would imply that Ψ_μ is a submersion. This is one reason why manifolds of finitely often differentiable mappings enter the picture here: The same statements as in the C^∞-case can be proven and these manifolds turn out to be Banach manifolds.

E.3.5 Let (M,g) be a Riemannian manifold with metric derivative ∇. Show that for vector fields $X,Y,Z \in \mathcal{V}(M)$, the following formula holds:

$$g(Z, \operatorname{grad} g(X,Y)) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y).$$

Then deduce that this implies $g(\nabla_X Y, Y) = \frac{1}{2} g(X, \operatorname{grad} g(Y,Y))$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Appendix F
Solutions to Selected Exercises

Chapter 1

Exercise 1.1.1 *Let $0 < p < 1$. We prove that the curve $\beta\colon [0,1] \to L^p[0,1]$, $\beta(t) := [\mathbf{1}_{[0,t[}]$ is an injective C^1-curve with $\beta'(t) = 0$, for all $t \in [0,1]$.*

Let us show that for every $x \in]0,1[$ the derivative of β vanishes. Then the claim follows by continuity also for the boundary points. Consider for h small the differential quotient $[h^{-1}(\beta(x+h) - \beta(x))] = h^{-1}[\mathbf{1}_{[x,x+h[}]$ converges to 0 with respect to the L^p-metric:

$$\begin{aligned} d([h^{-1}(\beta(x+h) - \beta(x))],[0]) &= \int_0^1 |h^{-1}\mathbf{1}_{[x,x+h[}(s)|^p \mathrm{d}s \\ &= |h|^{-p} \int_0^1 \mathbf{1}_{[x,x+h[}(s)\mathrm{d}s = |h|^{1-p}. \end{aligned}$$

Taking the limit $h \to 0$, we see that the derivative must be 0 and, in particular, β is a C^1-function. Now let $x < y$. Then $\beta(y) - \beta(x) = [\mathbf{1}_{[x,y[}] \neq [0]$, so β is injective.

Exercise 1.2.2(c) *We will show that $D\colon C^\infty([0,1],\mathbb{R}) \to C^\infty([0,1],\mathbb{R})$, $c \mapsto c'$ is continuous linear.*

Clearly the differential operator is linear with respect to pointwise addition of functions. Now let 0 be the constant 0-function. Then linearity of D together with Lemma A.5 implies that D will be continuous if the preimage of every 0-neighbourhood $U \subseteq C^\infty([0,1],\mathbb{R})$ is a 0-neighbourhood. Thus we pick an open 0-neighbourhood U. Shrinking U, we may assume that U is a ball $B_r^n(0)$ of radius $r > 0$ for the seminorm $\|\cdot\|_n$ where we have chosen suitable $n \in \mathbb{N}_0$ (and have exploited that these seminorms form a fundamental system by Example A.14). Now as the kth derivative of a function coincides with the $(k-1)$th derivative of its derivative, we observe that $D(B_{n+1}^r(0)) \subseteq B_n^r(0)$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

In other words, $B^r_{n+1}(0) \subseteq D^{-1}(U)$ and the preimage is a 0-neighbourhood. We deduce that D is continuous.

Exercise 1.3.3. Schwarz' theorem *We will show that for a C^k-map $f\colon E \supseteq U \to F$ and $x \in U$ the map $d^r f(x;\cdot)\colon E^r \to F$ is symmetric for every $2 \leq r \leq k$ (and $r < \infty$).*

Remark: There are several possibilities to prove this, for example, it suffices to show that the directional derivatives D_{v_1} and D_{v_2} commute for all $v_1, v_2 \in E$ (then the general case follows from $d^r f(x; v_1, \ldots, v_r) = D_{v_r} \cdots D_{v_1} f(x)$). However, we will reduce the problem to the well-known finite-dimensional case.

Use the Hahn–Banach theorem: It suffices to prove that $d^r(\lambda \circ f)(x;\cdot) = \lambda(d^r f(x;\cdot))$ is symmetric for every continuous linear functional λ. Hence without loss of generality $F = \mathbb{R}$. Now pick $2 \leq r \leq k$ $v_1, \ldots, v_r \in E$. Then there is $\varepsilon > 0$ such that $x + \sum_{i=1}^r t_i v_i$ is contained in U for all $|t_i| < \varepsilon$. Thus we can define the auxiliary function

$$h\colon \mathbb{R}^r \supseteq]-\varepsilon, \varepsilon[^r \to \mathbb{R}, \quad (t_1, \ldots, t_r) \mapsto f(x + t_1 v_1 + \cdots + t_r v_r).$$

By the chain rule h is C^k. Note that by the finite-dimensional version of Schwarz' theorem the partial derivatives of h commute. Now the statement for f follows from the chain rule and the observation (see Remark 1.15) that

$$\frac{\partial^r}{\partial_{t_1} \cdots \partial_{t_r}}\bigg|_{t_1, \ldots, t_r = 0} h(t_1, \ldots, t_r) = d^r f(x; v_1, \ldots, v_r).$$

Exercise 1.3.5 *For a C^2 map f and C^1 maps g, h we derive a formula for the derivative of $\phi := df \circ (g, h)$.*

This is an exercise in applying the chain rule and the rule on partial differentials (Proposition 1.20):

$$\begin{aligned} d\phi(x; y) &= d(df \circ (g,h))(x; y) \\ &= (d_1 df)(g(x), h(x); dg(x; y)) + (d_2 df)(g(x), h(x); dh(x; y)) \\ &= d^2 f(g(x); h(x), dg(x; y)) + df(g(x); dh(x; y)). \end{aligned}$$

Here we have used the fact that the derivative of df with respect to the first component is just $d^2 f$ and that $df(g(x);\cdot)$ is continuous linear.

Exercise 1.5.1 *We construct charts turning* $\operatorname{graph}(f) := \{(m, f(m)) \mid m \in M\}$ *into a split C^r submanifold of $M \times N$ for a C^r-function f.*

It suffices to construct submanifold charts for every point $(m, f(m))$. To this end, pick (U_φ, φ) a chart of M and (U_ψ, ψ) a chart of N with $m \in U_\varphi$, $f(m) \in U_\psi$. Assume that $\varphi \times \psi$ is a mapping into the locally convex space

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$E \times F$. We will now construct a chart around $(m, f(m))$ mapping all elements in $\operatorname{graph}(f) \cap U_\varphi \times U_\psi$ to the (complemented) subspace $E \times \{0\} \subseteq E \times F$. Since the vector space operations are continuous (bi)linear, they are smooth in the Bastiani sense and we obtain a C^r-map for $M \times N$ via

$$\kappa\colon U_\varphi \times U_\psi \to E \times F, \quad (m,n) \mapsto (\varphi(m), \psi(n) - \psi(f(m))).$$

This mapping is a C^r-diffeomorphism as its inverse is given by the formula $\kappa^{-1}(x,y) := (\varphi^{-1}(x), \psi^{-1}(y + \psi(f(\varphi^{-1}(x)))))$ (we leave it as an exercise to show that this mapping is well defined on an open subset of $E \times F$). Thus κ is a chart for $M \times N$. By construction $\kappa(m, f(m)) = (\varphi(m), \psi(f(m)) - \psi(f(m))) = (\varphi(m), 0)$. Thus κ is a submanifold chart for the graph.

Exercise 1.5.4 *We prove that a locally compact manifold M is necessarily finite dimensional* (note that the exercise asks for compact manifolds, but the argument only requires local compactness).

Let $\varphi\colon U_\varphi \to V_\varphi \subseteq E$ be a chart for M. Since M is locally compact, there exists a compact neighbourhood C of $x \in U_\varphi$ such that $C \subseteq U_\varphi$. Then $\varphi(C) \subseteq E$ is a compact neighbourhood of $\varphi(x)$. Since translations are homeomorphisms in E, the translated set $\varphi(C) - \varphi(x) = \{y = m - \varphi(x) \mid m \in \varphi(C)\}$ is a compact 0-neighbourhood in E. Thus E is finite dimensional by Proposition A.3 and M is a finite-dimensional manifold.

Exercise 1.7.1 *We show that the composition $g \circ f$ of the submersions $f\colon M \to N$ and $g\colon N \to L$ is a submersion.*

The submersion property is local, whence we can restrict to chart neighbourhoods of submersion charts around $m \in U \subseteq M, f(m) \in V \subseteq N$ and $g(f(m)) \in W \subseteq L$ such that:

$$\begin{array}{ccccccc}
U & \xrightarrow{f} & V & = & V & \xrightarrow{g} & W \\
\downarrow{\scriptstyle\varphi} & & \downarrow{\scriptstyle\psi_1} & & \downarrow{\scriptstyle\psi_2} & & \downarrow{\scriptstyle\kappa} \\
F \times X & \xrightarrow{\mathrm{pr}_F} & F & \xrightarrow{\cong} & Y \times Z & \xrightarrow{\mathrm{pr}_Y} & Y.
\end{array}$$

Now via the typical insertion of charts:

$$\begin{aligned}
\kappa \circ g \circ f \circ \varphi^{-1} &= \kappa \circ g \circ \psi_2^{-1} \circ \psi_2 \circ \psi_1^{-1} \circ \psi_1 \circ f \circ \varphi^{-1} \\
&= \mathrm{pr}_Y \circ \psi_2 \circ \psi_1^{-1} \circ \mathrm{pr}_F. \qquad \text{(F.1)}
\end{aligned}$$

Note that the change of charts $\psi_2 \circ \psi_1^{-1}$ is a diffeomorphism on its domain (which we will now call O). Shrinking U, we may assume that $\varphi(U) = O \times D$. We obtain a modified chart $\tilde{\varphi} := ((\psi_2 \circ \psi_1^{-1}) \times \mathrm{id}_W) \circ \varphi$. If we now insert $\tilde{\varphi}$ into (F.1) we see that $\kappa \circ g \circ f \circ \tilde{\varphi}^{-1} = \mathrm{pr}_Y\colon F \times X \cong Y \times Z \times X \to Y$. In other

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

words, we have constructed submersion charts for the composition which thus turns out to be a submersion.

Chapter 2

Exercise 2.2.1 *We assume that the exponential law holds for all relevant function spaces in this exercise! If $f \in C^\infty(U \times M, O)$ and $p \in C^\infty(O,N)$, we will prove that the pushforward satisfies $p_* \circ (f^\vee) = (p \circ f)^\vee$.*

Pick $(u,m) \in U \times M$. Then

$$\begin{aligned} p_* \circ (f^\vee)(u)(m) &= p_*(f^\vee(u))(m) = (p \circ f(u,\cdot))(m) = p(f(u,m)) \\ &= (p \circ f)^\vee(u)(m) \end{aligned}$$

and this proves the assertion. (Why did we need to assume that the exponential law holds if the calculation does not use it?)

Exercise 2.3.1 *Let $h\colon L \to K$ be a smooth map. Assume that $C^\infty(K,M)$ and $C^\infty(L,M)$ are canonical manifolds. We prove that*

(a) *the pullback $h^*\colon C^\infty(K,M) \to C^\infty(L,M)$, $f \mapsto f \circ h$ is smooth;*
(b) *if K, L are compact and M admits a local addition, then $TC^\infty(K,M) \cong C^\infty(K,TM)$ (see C.12). This identifies $T(h^*)$ with $h^*\colon C^\infty(K,TM) \to C^\infty(L,TM)$.*

(a) The pullback is a partial map of the full composition, hence smooth by Proposition 2.23. Alternatively, smoothness follows directly from the exponential law, as h^* is smooth if and only if the adjoint $(h^*)^\wedge\colon C^\infty(K,M) \times L \to M$, $(f,\ell) \mapsto f(h(\ell)) = \mathrm{ev}(f, h(\ell))$ is smooth. Since $C^\infty(K,M)$ is canonical, Lemma 2.16 shows that the evaluation is smooth. Now smoothness of the adjoint follows, since ev and h are smooth.

(b) We only need the assumptions to identify $TC^\infty(K,M) \cong C^\infty(K,TM)$. To compute the tangent, we pick $c\colon]-\varepsilon,\varepsilon[\to C^\infty(K,M)$ smooth with $c(0) = f$ and $\dot c(0) = V_f$. Under the identification we can interpret $V_f(x) = \frac{\partial}{\partial t}\big|_{t=0} c^\wedge(t,x)$ as a function $K \to TM$. Now

$$Th^*(V_f)(x) = \frac{\partial}{\partial t}\bigg|_{t=0} h^*(c^\wedge(t,x)) = \frac{\partial}{\partial t}\bigg|_{t=0} c^\wedge(t,h(x)) = V_f(h(x)) = h^*(V_f).$$

Thus we have identified the tangent map as $h^*\colon C^\infty(K,TM) \to C^\infty(L,TM)$.

Exercise 2.3.4 *For K a compact manifold and M a manifold with local addition, we endow $C^\infty(K,M)$ with its canonical manifold structure and compute the tangent map of the evaluation* $\mathrm{ev}\colon C^\infty(K,M) \times K \to M$.

We apply the rule on partial differentials for manifolds (Exercise 1.6.1):

$$T_{(\varphi,k)}\,\mathrm{ev}(v_\varphi, v_k) = T_\varphi\,\mathrm{ev}(\cdot,k)(v_\varphi) + T_k\,\mathrm{ev}(\varphi,\cdot)(v_k).$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

To evaluate the first term, pick a curve $c\colon]-\varepsilon,\varepsilon[\to C^\infty(K,M)$ with $c(0)=\varphi$ and $\dot c(0)=v_\varphi$. If we identify $TC^\infty(K,M)\cong C^\infty(K,TM)$ via C.12 we can interpret v_φ as the smooth mapping $\frac{\partial}{\partial t}\big|_{t=0} c^\wedge\colon K\to TM$. Then we compute

$$T_\varphi \operatorname{ev}(\cdot,k)(v_\varphi)=\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\operatorname{ev}(c(t),k)=\frac{\partial}{\partial t}\bigg|_{t=0} c^\wedge(t,k)=v_\varphi(k)=\operatorname{ev}_k(v_\varphi).$$

Here $\operatorname{ev}_k\colon C^\infty(K,TM)\to TM$ is the evaluation in k and we have exploited the exponential law in the computation. For the second term in the sum, it is immediately clear that we get $T_k\varphi(v_k)$. Thus we get as a formula for the tangent mapping,

$$T\operatorname{ev}(v_\varphi,v_k)=\operatorname{ev}_k(v_\varphi)+T\varphi(v_k).$$

Chapter 3

Exercise 3.1.5 *Let $H=\{(x_n)_{n\in\mathbb{N}}\in\ell^2 \mid x_n\in\frac{1}{n}\mathbb{Z}, n\in\mathbb{N}\}$. Then we prove that*

(a) *Every 0-neighbourhood in the subspace topology of H contains at least one non-zero element.*
(b) *There is no 0-neighbourhood in H which contains a continuous path connecting 0 with a non-zero element. Thus H is not a (sub)manifold.*

(a) It suffices to consider the intersection of norm balls with H, and in particular, we only need to find such elements in $B_{1/m}(0)\cap H$ for $m\in\mathbb{N}$. However, for such a ball, it is clear that the sequence $x_n^m=0$ if $m\neq n$ and $x_m^m=1/(2m)$ is contained in the intersection.

(b) Assume that there is a 0-neighbourhood in H which is path-connected. Then it contains an element $(x_k)_{k\in\mathbb{N}}\neq 0$. Pick $\ell\in\mathbb{N}$ with $x_\ell\neq 0$. If $c\colon[0,1]\to H$ is a continuous path connecting 0 and $(x_k)_{k\in\mathbb{N}}$, then $\pi_\ell\circ c$ is a continuous path in $\mathbb{R}$ connecting 0 and $x_\ell\neq 0$. Since the path c takes its values in H, $\pi_\ell\circ c$ can take only values in a discrete subset of $\mathbb{R}$. Contradiction! Thus there is no 0-neighbourhood of H in the subspace topology which is path-connected and H is therefore not locally homeomorphic to a locally convex space. We conclude that it cannot be a (sub-)manifold of ℓ^2.

Exercise 3.3.7 *(Mini Lie–Palais) Every Lie algebra morphism $\phi\colon\mathfrak{g}\to\mathcal{V}(M)$ from a finite-dimensional Lie algebra $\mathfrak{g}$ to the Lie algebra of vector fields of a compact manifold M (with the negative of the usual bracket) gives rise to a Lie group action $G\times M\to M$ (with $\mathbf{L}(G)=\mathfrak{g}$).*

Note first that since $\mathfrak{g}$ is finite dimensional and ϕ is linear, ϕ is automatically continuous, hence a morphism of locally convex Lie algebras. By Lie's third theorem (Hilgert and Neeb, 2012, Theorem 9.4.11) there exists a connected,

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

simply connected Lie group G such that $\mathbf{L}(G) = \mathfrak{g}$. Now G is finite dimensional and thus regular and $\mathrm{Diff}(M)$ is regular by Example 3.36. Hence we can apply Lie's second theorem for regular Lie groups, Proposition E.14, to integrate ϕ to a Lie group morphism $\Phi\colon G \to \mathrm{Diff}(M)$, that is, $\mathbf{L}(\Phi) = \phi$. Exploiting that $\mathrm{Diff}(M) \subseteq C^\infty(M,M)$ and $C^\infty(M,M)$ is a canonical manifold, Lemma C.11, the adjoint map $\Phi^\wedge\colon G \times M \to M$ is a Lie group action (naturally induced by ϕ).

Chapter 5

Exercise 5.1.1(c) *We check the details of Proposition 5.3 and show that the second derivative of the vertical part of S_* is given by the pushforward of the second derivative of the vertical part of S.*

We already saw in the proposition that we can instead compute the partial derivative (with respect to the first variable) of

$$(S_*)^\wedge\colon C^\infty(\mathbb{S}^1,TM)\times\mathbb{S}_1 \to T^2M, \quad (h,\theta)\mapsto S(h(\theta)).$$

Pick $\theta\in\mathbb{S}^1$ and $h\in C^\infty(\mathbb{S}^1,TM)$ together with a chart (U,φ) of M such that $h(\theta)\in TU \cong U\times E$ (for E a locally convex space). By continuity, there is a compact set L such that $h\in\lfloor L,TU\rfloor$ and $\theta\in L^\circ$. For a curve $c\colon\,]-\varepsilon,\varepsilon[\to\lfloor L,TU\rfloor$, apply again the exponential law, Theorem 2.12, to see that $c(t)|_{L^\circ} = T\varphi^{-1}c_E(t)$ for some smooth $c_E\colon\,]-\varepsilon,\varepsilon[\to C^\infty(L^\circ,U\times E)$. Plugging this into $(S_*)^\wedge$ we take derivatives to obtain

$$\begin{aligned}
&T^2\varphi\circ(S_*)^\wedge(T\varphi^{-1})_*\circ c_E(t),\theta)\\
&\quad= T^2\varphi S(T\varphi^{-1}(c_E(t)(\theta))\\
&\quad= (\pi_M\circ h(\theta),c_E(t)(\theta),c_E(t)(\theta),S_{U,2}(h(\theta),c_E(t)(\theta))).
\end{aligned}$$

Here $S_{U,2}$ is the non-trivial vector part of the spray S. We conclude that after projecting onto the fourth component and after fixing the parameter θ, the second derivative of the vertical part of S_* can be identified with the pushforward of the second derivative of $S_{U,2}$. This proves that B_{S_*} is the pushforward of B.

Exercise 5.1.2 *We establish the formula (5.5): $\nabla_{\dot c}\alpha(\cdot)(x) = \nabla^g_{\dot c^\wedge(\cdot,x)}\alpha^\wedge(\cdot,x)$ for all $x\in\mathbb{S}^1$.*

The trick is to avoid at all costs working in charts of the manifold of mappings $C^\infty(\mathbb{S}^1,M)$. However, the object $\nabla_{\dot c}\alpha$ is defined via the local formula (4.13). Taking a look at the local formula for the connector (4.11), we see that $\nabla_{\dot c}\alpha = K_*(\dot\alpha)$, where K_* is the connector associated to the covariant derivative ∇. As already shown in the notation (and proved in Proposition 5.7), the

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

connector of ∇ is the pushforward of the connector K associated to the Riemannian metric g on M (and thus also to the covariant derivative ∇^g). Setting in now these relations, we obtain the desired identity

$$\nabla_{\dot{c}}\alpha(t)(x) = K_*(\dot{\alpha}(t))(x) = (K \circ \alpha^\wedge)(t,x) = \nabla_{c^\wedge(t,x)}\alpha^\wedge(t,x).$$

Exercise 5.2.3 *The elastic metric* (5.9) *is invariant under reparametrisations with elements* φ *in* $\mathrm{Diff}(\mathbb{S}^1)$ *which satisfy* $T_\theta\varphi(1) > 0$ *for all* $\theta \in \mathbb{S}^1$.

We have seen in Proposition 5.17 that the elastic metric is the pullback of the L^2-metric via the SRVT. In Exercise 5.2.1(c) we have seen that for a diffeomorphism φ in $\mathrm{Diff}(\mathbb{S}^1)$ which satisfies $T_\theta\varphi(1) > 0$, for all $\theta \in \mathbb{S}^1$ one has $\mathcal{R}(c \circ \varphi) = \dot{\varphi} \cdot \mathcal{R}(c) \circ \varphi$. Plugging this in the L^2-inner product, we see that invariance follows from the usual transformation rule for integrals.

Chapter 6

Exercise 6.2.3 *We show that for a Banach Lie groupoid* $G \rightrightarrows M$ *the multiplication map* $\mathbf{m}\colon G \times_M G \to G$ *is a submersion*.

Remark: I do not know whether this proof (which was shared with me by D. M. Roberts (Adelaide)) or even the corresponding statement generalises beyond the Banach setting.

We exploit the following characterisation of submersions between Banach manifolds: The map $\mathbf{m}$ is a submersion if and only if it admits local sections, that is, for every $(g_1, g_2) \in G \times_M G$ there exists a smooth map $\varphi\colon U \to G \times_M G$ such that $g := \mathbf{m}(g_1, g_2) \in U$, $\varphi(g) = (g_1, g_2)$ and $\mathbf{m} \circ \varphi = \mathrm{id}_U$ (see Margalef-Roig and Domínguez, 1992, Proposition 4.1.13). Thus we fix g_1, g_2 and g as above and write $\varphi = (\varphi_1, \varphi_2) \in C^\infty(G, G \times G)$ (exploiting that $G \times_M G$ is a submanifold of the cartesian product, whence it suffices to obtain two smooth maps with values in G such that their combination takes values in the fibre-product). Now exploiting the groupoid structure, we observe that $g_2 = g_1^{-1} \cdot g$. Ignoring for a moment that multiplication is not globally defined, we see that for any smooth map φ_1 with $\varphi_1(g) = g_1$, the smoothness of the groupoid operations yields a smooth map φ_2 via

$$\varphi_2(x) = \varphi_1(x)^{-1} \cdot x = \mathbf{m}(\mathbf{i}(\varphi_1(x)), x). \tag{F.2}$$

Setting in $x = g$, we immediately see that $\varphi_2(g) = g_2$. However, to make the formula (F.2) well defined we need to require that $\mathbf{s}(\varphi_1(x)^{-1}) = \mathbf{t}(x)$. Since inversion intertwines source and target this yields $\mathbf{t} \circ \varphi_1 = \mathbf{t}$. We deduce that it suffices to construct a certain smooth map $\varphi_1\colon U \to G$ on some neighbourhood U of g with $\mathbf{t} \circ \varphi_1 = \mathbf{t}$ and $\varphi_1(g) = g_1$.

Set $y := \mathbf{t}(g)$ and observe that $\mathbf{t}(g) = \mathbf{t}(g_1 g_2) = \mathbf{t}(g_1)$. Since $\mathbf{t}$ is a submersion, there is an open neighbourhood O_y of y together with a smooth section

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

$\sigma\colon O_y \to G$ such that $\sigma(y) = g_1$ and $\mathbf{t}\circ\sigma = \mathrm{id}_{O_y}$. We can now choose an open neighbourhood $g \in U$ such that $\mathbf{t}(U) \subseteq O_y$ and define $\varphi_1\colon U \to G$, $x \mapsto \sigma \circ \mathbf{t}(x)$. Then φ_1 is smooth and satisfies $\varphi_1(g) = g_1$ and $\mathbf{t}\circ\varphi_1 = \mathbf{t}$. We conclude that $\mathbf{m}$ admits local sections and is thus a submersion.

Exercise 6.2.4 *We check that the gauge groupoid associated to a principal G-bundle (E,p,M,F) is a Lie groupoid if E, M and G are Banach manifolds.*

(a) We begin with the construction of charts for $(E\times E)/G$. Let $(U_i)_{i\in I}$ be an open cover of M such that there exist smooth local sections $\sigma_i\colon U_i \to E$ of p. This yields an atlas $(U_i,\kappa_i)_{i\in I}$ of local trivialisations of the bundle $p\colon E\to M$ which are given by

$$\kappa_i\colon p^{-1}(U_i) \to U_i\times G, \quad x\mapsto (p(x),\mathrm{d}(\sigma_i(p(x)),x)),$$

with $\mathrm{d}\colon E\times_M E\to G$, $(x,y)\mapsto x^{-1}\cdot y$. Here we use $x^{-1}\cdot y$ as the suggestive notation for the unique element $g\in G$ that satisfies $x\cdot g = y$.
The local trivialisations commute with the right G-action on E since

$$\kappa_i(x\cdot g) = (p(x\cdot g),\mathrm{d}(\sigma_i(p(x\cdot g)),x\cdot g) = (p(x),\mathrm{d}(\sigma_i(p(x)),x))\cdot g.$$

In particular, the trivialisations descend to manifold charts for the arrow manifold of the gauge groupoid:

$$\begin{aligned} &K_{ij}\colon (p^{-1}(U_i)\times p^{-1}(U_j))/G \to U_i\times U_j\times G,\\ &[x_1,x_2]\mapsto (p(x_1),p(x_2),\mathrm{d}(\sigma_i(p(x_1)),x_1)\mathrm{d}(\sigma_j(p(x_2)),x_2)^{-1}).\end{aligned}$$

To see that the projection is a submersion, it suffices to prove this locally in charts. In the trivialisations and the charts, the quotient becomes the map

$$(U_i\times G)\times(U_j\times G)\to U_i\times U_j\times G, \quad \big((u_i,g_i),(u_j,g_j)\big)\mapsto (u_i,u_j,g_ig_j^{-1}).$$

While we have the identity in the u components, the G component is the composition of inversion in the second component with the Lie group multiplication. Inversion in a Lie group is a diffeomorphism, while the multiplication in the Banach Lie group G is a submersion by Exercise 6.2.3. Now the composition of submersions is a submersion by Exercise 1.7.1 and by Exercise 1.7.2, the quotient map is a submersion.

(b) Smoothness of the mappings follows from Exercise 1.7.6 by composing them with the submersion $q\colon E\times E\to (E\times E)/G$ and observing that the resulting mappings are smooth on $E\times E$. In particular, $\mathbf{s}\circ q = p$ is a surjective submersion, whence by Margalef-Roig and Domínguez (1992, Proposition 4.1.5), $\mathbf{s}$ is a surjective submersion.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Chapter 7

Exercise 7.3.1 *Let (M,g) be a Riemannian metric with volume form μ. We show that the L^2-metric is a right invariant Riemannian metric on* $\mathrm{Diff}_\mu(M)$.

We have already seen (for $M = \mathbb{S}^1$ but the general case is similar) that the L^2-inner product $\langle X,Y\rangle_{L^2} := \int_M g_m(X(m),Y(m))\mathrm{d}\mu(m)$ is a continuous inner product on $\mathcal{V}(M)$. From the formula for the tangent of the right multiplication in $\mathrm{Diff}(M)$ (see 2.22), we see that the right-invariant metric induced by the L^2-inner product is given by

$$\langle X\circ\varphi, Y\circ\varphi\rangle_\varphi = \langle X\circ\varphi\circ\varphi^{-1}, Y\circ\varphi\circ\varphi^{-1}\rangle_{L^2} = \langle X,Y\rangle_{L^2}.$$

Let us assume now that φ is a volume-preserving diffeomorphism. Then by diffeomorphism invariance of the integral (see e.g. Lee, 2013, Proposition 16.6), we derive now that

$$\begin{aligned}\langle X\circ\varphi, Y\circ\varphi\rangle_\varphi = \langle X,Y\rangle_{L^2} &= \int_M \varphi^* g(X,Y)\mathrm{d}\mu \\ &= \int_M g_{\varphi(m)}(X\circ\varphi(m), Y\circ\varphi(m))\mathrm{d}\mu(m).\end{aligned}$$

Thus for every volume-preserving diffeomorphism, the L^2-metric coincides with the right-invariant Riemannian metric induced by the L^2-inner product on the vector fields. We conclude that the L^2-metric is right invariant on the subgroup $\mathrm{Diff}_\mu(M)$.

Exercise 7.3.1 *For (M,g) a compact Riemannian manifold consider the L^2-metric g^{L^2} (7.8) on* $\mathrm{Diff}(M)$. *Let S be the metric spray of g and K the associated connector. (a) Then S_* and K_* define a spray and a connector on* $\mathrm{Diff}(M)$. *Moreover, by (b)–(c) $\nabla^{L^2}_X Y := K_*\circ TY\circ X$ defines the metric derivative on g^{L^2}.*

(a) As $\mathrm{Diff}(M) \subseteq C^\infty(M,M)$ and $C^\infty(M,M)$ is a canonical manifold, the pushforwards are smooth. That they form a spray and a connector follows directly from the identification of $T^k\,\mathrm{Diff}(M) \subseteq T^kC^\infty(M,M) \cong C^\infty(M,T^kM)$.

(b)–(c) Details for this proof are recorded in Ebin and Marsden (1970, Proof of Theorem 9.1).

Chapter 8

Exercise 8.2.4 *Let $N \in \mathbb{N}\cup\{\infty\}, d \in \mathbb{N}$ and $X\colon [0,1]\to\mathbb{R}^d$ be a smooth path. Define $DX\colon [0,1]\to\mathcal{T}^N(\mathbb{R}^d)$, $t\mapsto(0,X'_t,0,\ldots)$. We will then:*

(a) *Show that* $\frac{\mathrm{d}}{\mathrm{d}t}S_N(X)_{s,t} = S_N(X)_{s,t}\otimes(DX)_t, s<t\le 1, S_N(X)_{s,s} = \mathbf{1}$.
(b) *Establish Chen's relation*

$$S_N(X)_{s,t} = S_N(X)_{s,u}\otimes S_N(X)_{u,t}, \qquad 0\le s\le u\le t\le 1.$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

(c) *Thus* $S_N(X)_{s,t} = S_N(X)_{0,s}^{-1} \otimes S_N(X)_{0,t}$ *holds*.

First note that the claims will follow for $N \in \mathbb{N} \cup \{\infty\}$ if we can establish the identity for the projection to every finite degree $k \in \mathbb{N}_0$.

(a) For $s = 0$ the claim is just (8.7). For arbitrary s the claim follows from inspecting the Volterra series. Namely, projecting to homogeneous elements of degree $k \geq 1$, the component of $S_N(X)_{s,t}$ is

$$\begin{aligned}
&\int_s^t \int_s^{r_{k-1}} \cdots \int_s^{r_1} \mathrm{d}X_{r_1} \otimes \cdots \otimes \mathrm{d}X_{r_k} \\
&\quad = \int_s^t \left(\int_s^{r_{k-1}} \cdots \int_s^{r_1} \mathrm{d}X_{r_1} \otimes \cdots \otimes \mathrm{d}X_{r_{k-1}} \right) \otimes \mathrm{d}X_{r_k} \\
&\quad = \int_s^t \pi_{k-1}^N \left(S_N(X)_{s,r_k} \right) \otimes \mathrm{d}X_{r_k}.
\end{aligned}$$

In other words, the signature satisfies the integral equation $S_N(X)_{s,t} = 1 + \int_s^t S_N(X)_{s,r} \otimes \mathrm{d}X_r$ in $\mathcal{T}^N(\mathbb{R}^d)$, whence it solves the desired ODE.

(b) We proceed by induction on k. Note that the identity is trivially true for $k = 0$ since it reads $1 = 1 \cdot 1$. Assume now that we have established the claim now for every $s < u < t \in [0,1]$ and $\ell \leq k$, and so $S_k(X)_{s,t} = S_k(X)_{s,u} \otimes S_k(X)_{u,t}$. We work now in the truncated tensor algebra $\mathcal{T}^{k+1}(\mathbb{R}^d)$ (and note that the following identities hold precisely by truncating after degree $k+1$):

$$S_{k+1}(X)_{s,u} = 1 + \int_s^u S_{k+1}(X)_{s,r} \otimes \mathrm{d}X_r = \int_s^u S_k(X)_{s,r} \otimes \mathrm{d}X_r,$$

$$S_{k+1}(X)_{s,u} \otimes \int_u^t S_N(X)_{u,r} \otimes \mathrm{d}X_r = S_k(X)_{s,u} \otimes \int_u^t S_k(X)_{u,r} \otimes \mathrm{d}X_r.$$

Applying the induction hypothesis to split the $S_k(X)$ for $s < u < r < t$, we obtain

$$\begin{aligned}
S_{k+1}(X)_{s,t} &= 1 + \int_s^u S_k(X)_{s,r} \mathrm{d}X_r + \int_u^t S_k(X)_{s,u} \otimes S_k(X)_{u,r} \otimes \mathrm{d}X_r \\
&= S_{k+1}(X)_{s,u} + S_{k+1}(X)_{s,u} \otimes \left(\int_u^t S_k(X)_{t,r} \otimes \mathrm{d}X_r \right) \\
&= S_{k+1}(X)_{s,u} \otimes (1 + (S_{k+1}(X)_{u,t} - 1)) \\
&= S_{k+1}(X)_{s,u} \otimes S_{k+1}(X)_{u,t}.
\end{aligned}$$

(c) Multiplying Chen's relation for $S_N(X)_{0,t}$ from the left with the inverse of $S_N(X)_{0,s}$ immediately yields the desired identity.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Appendix A

Exercise A.6.2(a) *We show that the box topology turns* $E := \{(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \text{ almost all } x_n = 0\}$ *into a locally convex space and makes the right shift continuous.*

Note that we have $\lambda(\prod_n U_n) + (\prod_n V_n) = \prod_{n\in\mathbb{N}}(\lambda U_n + V_n)$ for $U_n, V_n \Subset \mathbb{R}$ and $\lambda \in \mathbb{R}$. If $(x_n) + (y_n) \in \prod_n U_n$ we exploit that $\mathbb{R}$ is a topological vector space to find for every $n \in \mathbb{N}$ $x_n \in A_n \Subset \mathbb{R}$ and $y_n \in B_n \Subset \mathbb{R}$ with $A_n + B_n \subseteq U_n$. We conclude that $(\prod_n A_n) + (\prod_n B_n) \subseteq (\prod_n U_n)$ and vector addition is continuous. For scalar multiplication we note that if $\lambda \cdot (x_n) \in \prod_n U_n$ we can exploit that $x_n \neq 0$ for only finitely many n, to construct an open λ-neighbourhood $\lambda \in O \Subset \mathbb{R}$ together with $(x_n) \in \prod_n V_n$ such that $O \cdot \prod_n V_n \subseteq \prod_n U_n$. Hence scalar multiplication is continuous and E is a TVS.

Now let $\prod_n U_n$ be a 0-neighbourhood. This implies that every U_n is a 0-neighbourhood in $\mathbb{R}$. Since $\mathbb{R}$ is locally convex, for every n there is a convex 0-neighbourhood $C_n \subseteq U_n$. Then $\prod_n C_n \subseteq \prod_n U_n$ is a convex 0-neighbourhood and thus E is locally convex. The right shift is continuous as $R^{-1}(\prod_n U_n) = \prod_n U_{n+1}$ if $0 \in U_1$ and $\emptyset$ otherwise.

Appendix B

Exercise B.2.4 *We show that the set* $\Omega' := \{f \in C(K,Y) \mid \operatorname{graph}(f) \subseteq \Omega\}$ *is open in the compact-open topology if* K *is compact and* $\Omega \Subset K \times Y$.

Let us show that Ω' is a neighbourhood for each $f \in \Omega'$. Since Ω is open in $K \times Y$ with the product topology, we find for each $x \in K$ open subsets $U_x \Subset K, V_x \Subset Y$ with $(x, f(x)) \in U_x \times V_x \Subset \Omega$. Shrinking the U_x, we may also assume that $\overline{U}_x \times V_x \subseteq \Omega$ and $f(\overline{U}_x) \subseteq V_x$.

By compactness $\overline{U}_x$ is compact and we can cover K with finitely many of the U_x, say, $K = \bigcup_{1\leq k\leq n} U_{x_k}$. Then by construction $N_f := \bigcap_{1\leq k\leq n} \lfloor \overline{U}_{x_k}, V_{x_k} \rfloor$ is open in the compact-open topology and $f \in N_f$. Moreover, if $h \in N_f$, then we have for $x \in U_{x_k}$ that $(x, h(x)) \in U_x \times V_x \subseteq \Omega$. Since the U_x cover K, we deduce that $h \in \Omega'$ and thus $N_f \subseteq \Omega'$.

Appendix C

Exercise C.2.2 We show that the pullback bundle $f^*(E)$ is a split submanifold of $K \times E$. The idea is similar to the proof that the graph of a smooth function is a split submanifold. The proof is essentially Lemma 1.60: By definition of a fibre bundle $p\colon E \to M$ is a submersion. In the notation of that lemma we have $f^*(E) = K \times_M E$ is a split submanifold of $K \times E$ for each $f \in C^\infty(K, M)$.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Appendix E

Exercise E.3.5 *For a Riemannian manifold* (M,g) *with metric derivative* ∇*, we establish the formula*

$$g(Z,\operatorname{grad} g(X,Y)) = g(\nabla_Z X,Y) + g(X,\nabla_Z Y) \text{ for } X,Y,Z \in \mathcal{V}(M).$$

From the definition of the gradient and the compatibility of the metric derivative with g we deduce that

$$g(Z,\operatorname{grad} g(X,Y)) = dg(X,Y)(\cdot;Z) = Z\,.\,g(X,Y) = g(\nabla_Z X,Y) + g(X,\nabla_Z Y).$$

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

References

Abbati, M. C., Cirelli, R., Mania', A., and Michor, P. 1989. The Lie group of automorphisms of a principal bundle. *J. Geom. Phys.*, **6**(2), 215–235.

Abraham, R., Marsden, J. E., and Ratiu, T. 1988. *Manifolds, Tensor Analysis, and Applications*. Second ed. Applied Mathematical Sciences, vol. 75. Springer, New York.

Agrachev, A. A., and Caponigro, M. 2009. Controllability on the group of diffeomorphisms. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, **26**(6), 2503–2509.

Alzaareer, H., and Schmeding, A. 2015. Differentiable mappings on products with different degrees of differentiability in the two factors. *Expo. Math.*, **33**(2), 184–222.

Amiri, H., Glöckner, H., and Schmeding, A. 2020. Lie groupoids of mappings taking values in a Lie groupoid. *Arch. Math.*, **56**(5), 307–356.

Amiri, H., and Schmeding, A. 2019. A differentiable monoid of smooth maps on Lie groupoids. *J. Lie Theory*, **29**(4), 1167–1192.

Amiri, H., and Schmeding, A. 2019. Linking Lie groupoid representations and representations of infinite-dimensional Lie groups. *Ann. Global Anal. Geom.*, **55**(4), 749–775.

Arnold, V. 1966. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier (Grenoble)*, **16**(1), 319–361.

Atkin, C. J. 1975. The Hopf-Rinow theorem is false in infinite dimensions. *Bull. London Math. Soc.*, **7**(3), 261–266.

Atkin, C. J. 1997. Geodesic and metric completeness in infinite dimensions. *Hokkaido Math. J.*, **26**(1), 1–61.

Baez, J. C. 1997. An introduction to *n*-categories. Pages 1–33 of: *Category Theory and Computer Science. 7th International Conference, CTCS '97, Santa Margherita Ligure, Italy, September 4–6, 1997. Proceedings*. Springer, Berlin.

Banyaga, A. 1988. On isomorphic classical diffeomorphism groups. II. *J. Differ. Geom.*, **28**(1), 23–35.

Bastiani, A. 1964. Applications différentiables et variétés différentiables de dimension infinie. *J. Analyse Math.*, **13**, 1–114.

Bauer, M., and Modin, K. 2020. Semi-invariant Riemannian metrics in hydrodynamics. *Calc. Var. Partial Differ. Equ.*, **59**(2), 25. Id/No 65.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Bauer, M., Bruveris, M., and Michor, P. W. 2014. Homogeneous Sobolev metric of order one on diffeomorphism groups on real line. *J. Nonlinear Sci.*, **24**(5), 769–808.

Bauer, M., Bruveris, M., and Michor, P. W. 2014. Overview of the geometries of shape spaces and diffeomorphism groups. *J. Math. Imaging Vis.*, **50**(1–2), 60–97.

Baum, H. 2014. *Eichfeldtheorie. Eine Einführung in die Differentialgeometrie auf Faserbündeln*. Second revised ed. Springer Spektrum, Heidelberg.

Beggs, E. J. 1987. The de Rham complex on infinite-dimensional manifolds. *Quart. J. Math. Oxford II Ser.*, **38**(150), 131–154.

Beltiţă, D., Goliński, T., and Tumpach, A.-B. 2018. Queer Poisson brackets. *J. Geom. Phys.*, **132**, 358–362.

Beltiţă, D., Goliński, T., Jakimowicz, G., and Pelletier, F. 2019. Banach–Lie groupoids and generalized inversion. *J. Funct. Anal.*, **276**(5), 1528–1574.

Bertram, W., Glöckner, H., and Neeb, K.-H. 2004. Differential calculus over general base fields and rings. *Expo. Math.*, **22**(3), 213–282.

Bogfjellmo, G., and Schmeding, A. 2017. The Lie group structure of the Butcher group. *Found. Comput. Math.*, **17**(1), 127–159.

Bogfjellmo, G., and Schmeding, A. 2018. The geometry of characters of Hopf algebras. Pages 159–185 of: *Computation and Combinatorics in Dynamics, Stochastics and Control. The Abel Symposium, Rosendal, Norway, August 16–19, 2016. Selected papers*. Springer, Cham.

Bogfjellmo, G., Dahmen, R., and Schmeding, A. 2016. Character groups of Hopf algebras as infinite-dimensional Lie groups. *Ann. Inst. Fourier*, **66**(5), 2101–2155.

Boman, J. 1967. Differentiability of a function and of its compositions with functions of one variable. *Math. Scand.*, **20**, 249–268.

Bonic, R., and Frampton, J. 1966. Smooth functions on Banach manifolds. *J. Math. Mech.*, **15**, 877–898.

Bourbaki, N. 1998. *Lie groups and Lie algebras. Chapters 1–3*. Elements of Mathematics. Springer, Berlin. Translated from the French, reprint of the 1989 English translation.

Bridson, M. R., and Haefliger, A. 1999. *Metric Spaces of Non-Positive Curvature*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 319. Springer, Berlin.

Bröcker, T., and Tom Dieck, T. 1995. *Representations of Compact Lie Groups*. Graduate Texts in Mathematics, vol. 98. Springer, New York. Translated from the German manuscript, corrected reprint of the 1985 translation.

Bruveris, M. 2018. *The L^2-metric on $C^\infty(M,N)$*. arXiv:1804.00577

Bruveris, M. 2019. *Riemannian Geometry for Shape Analysis and Computational Anatomy*. Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore. Pages 15–44. www.worldscientific.com/doi/pdf/10.1142/9789811200137_0002

Celledoni, E., Eslitzbichler, M., and Schmeding, A. 2016. Shape analysis on Lie groups with applications in computer animation. *J. Geom. Mech.*, **8**(3), 273–304.

Celledoni, E., Eidnes, S., and Schmeding, A. 2018. Shape analysis on homogeneous spaces: a generalised SRVT framework. Pages 187–220 of: *Computation and Combinatorics in Dynamics, Stochastics and Control. The Abel Symposium, Rosendal, Norway, August 16–19, 2016. Selected Papers*. Springer, Cham.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Chevyrev, I., and Kormilitzin, A. 2016. *A Primer on the Signature Method in Machine Learning*. arXiv:1603.03788

Curry, C., Ebrahimi-Fard, K., Manchon, D., and Munthe-Kaas, H. Z. 2020. Planarly branched rough paths and rough differential equations on homogeneous spaces. *J. Differ. Equations*, **269**(11), 9740–9782.

Dahmen, R., and Schmeding, A. 2020. Lie groups of controlled characters of combinatorial Hopf algebras. *Ann. Inst. Henri Poincaré D, Comb. Phys. Interact. (AIHPD)*, **7**(3), 395–456.

del Hoyo, M., and Fernandes, R. L. 2018. Riemannian metrics on Lie groupoids. *J. Reine Angew. Math.*, **735**, 143–173.

do Carmo, M. P. 1992. *Riemannian Geometry*. Translated from the Portuguese by Francis Flaherty. Birkhäuser, Boston, MA.

Dobrowolski, T. 1995. Every infinite-dimensional Hilbert space is real-analytically isomorphic with its unit sphere. *J. Funct. Anal.*, **134**(2), 350–362.

Duistermaat, J. J., and Kolk, J. A. C. 2000. *Lie Groups*. Springer, Berlin.

Ebin, D. G. 2015. Groups of diffeomorphisms and fluid motion: reprise. Pages 99–105 of: *Geometry, Mechanics, and Dynamics. The Legacy of Jerry Marsden. Selected papers presented at a focus program, Fields Institute for Research in Mathematical Sciences, Toronto, Canada, July 2012*. Springer, New York.

Ebin, D. G., and Marsden, J. 1970. Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. of Math.*, **92**(2), 102–163.

Eells, Jr., J. 1966. A setting for global analysis. *Bull. Amer. Math. Soc.*, **72**, 751–807.

Eftekharinasab, K., and Petrusenko, V. 2020. Finslerian geodesics on Fréchet manifolds. *Bull. Transilv. Univ. Bras. III: Math. Inform. Phys.*, **13**(62), 129–152. doi:10.31926/but.mif.2020.13.62.1.11

Engelking, R. 1989. *General Topology*. Second ed. Sigma Series in Pure Mathematics, vol. 6. Heldermann Verlag, Berlin. Translated from the Polish by the author.

Filipkiewicz, R. P. 1982. Isomorphisms between diffeomorphism groups. *Ergodic Theory Dynam. Systems*, **2**(2), 159–171 (1983).

Folland, G. B., and Stein, E. M. 1982. *Hardy Spaces on Homogeneous Groups*. Vol. 28. Princeton University Press, Princeton, NJ.

Friz, P. K., and Hairer, M. 2020. *A Course on Rough Paths. With an Introduction to Regularity Structures*. Second ed. Springer, Cham.

Friz, P. K., and Victoir, N. B. 2010. *Multidimensional Stochastic Processes as Rough Paths. Theory and Applications*. Vol. 120. Cambridge University Press, Cambridge.

Frölicher, A., and Kriegl, A. 1988. *Linear Spaces and Differentiation Theory*. Pure and Applied Mathematics (New York). John Wiley & Sons, Chichester.

Gallot, S., Hulin, D., and Lafontaine, J. 2004. *Riemannian Geometry*. Third ed. Springer, Berlin.

Glöckner, H., and Neeb, K.-H. *Infinite-Dimensional Lie Groups*. in preparation.

Glöckner, H. 2002. Infinite-dimensional Lie groups without completeness restrictions. Pages 43–59 of: *Geometry and Analysis on Finite- and Infinite-Dimensional Lie Groups (Bedlewo, 2000)*. Banach Center Publications, vol. 55. Polish Acad. Sci. Inst. Math., Warsaw.

Glöckner, H. 2006a. Discontinuous non-linear mappings on locally convex direct limits. *Publ. Math. Debrecen*, **68**(1–2), 1–13.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Glöckner, H. 2006b. Implicit functions from topological vector spaces to Banach spaces. *Israel J. Math.*, **155**, 205–252.

Glöckner, H. 2007. *Implicit Functions from Topological Vector Spaces to Fréchet Spaces in the Presence of Metric Estimates*. arXiv:math/0612673v5

Glöckner, H. 2016. *Fundamentals of Submersions and Immersions between Infinite-Dimensional Manifolds*. arXiv:1502.05795v4

Glöckner, H., and Neeb, K.-H. 2012. When unit groups of continuous inverse algebras are regular Lie groups. *Studia Math.*, **211**(2), 95–109.

Glöckner, H., and Schmeding, A. 2022. *Manifolds of Mappings on Cartesian Products*. Ann Glob Anal Geom 61, 359–398. doi:10.1007/s10455-021-09816-y

Gray, W. S., Palmstrøm, M., and Schmeding, A. 2022. *Continuity of Formal Power Series Products in Nonlinear Control Theory*. Foundations of Computational Mathematics, doi:10.1007/s10208-022-09560-0

Grong, E., Markina, I., and Vasil'ev, A. 2015. Sub-Riemannian geometry on infinite-dimensional manifolds. *J. Geom. Anal.*, **25**(4), 2474–2515.

Gubinelli, M. 2010. Ramification of rough paths. *J. Differ. Equations*, **248**(4), 693–721.

Hamilton, R. S. 1982. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)*, **7**(1), 65–222.

Hilgert, J., and Neeb, K.-H. 2012. *Structure and Geometry of Lie Groups*. Springer Monographs in Mathematics. Springer, New York.

Hirsch, M. W. 1994. *Differential Topology*. *Graduate Texts in Mathematics*. Vol. 33. Springer, New York. Corrected reprint of the 1976 original.

Hjelle, E. O., and Schmeding, A. 2017. Strong topologies for spaces of smooth maps with infinite-dimensional target. *Expo. Math.*, **35**(1), 13–53.

Hofmann, K. H., and Neeb, K.-H. 2009. Pro-Lie groups which are infinite-dimensional Lie groups. *Math. Proc. Camb. Philos. Soc.*, **146**(2), 351–378.

Hofmann, K. H., and Morris, S. A. 2007. *The Lie Theory of Connected Pro-Lie Groups. A Structure Theory for Pro-Lie Algebras, Pro-Lie Groups, and Connected Locally Compact Groups*. Vol. 2. European Mathematical Society (EMS), Zürich.

Husemoller, D. H. 1993. *Fibre Bundles*. Third ed. Springer, Berlin.

Iglesias-Zemmour, P. 2013. *Diffeology*. Mathematical Surveys and Monographs, vol. 185. American Mathematical Society, Providence, RI.

Inci, H., Kappeler, T., and Topalov, P. 2013. *On the Regularity of the Composition of Diffeomorphisms*. American Mathematical Society, Providence, RI.

Jarchow, H. 1981. *Locally Convex Spaces*. B. G. Teubner, Stuttgart.

Keller, H. H. 1974. *Differential Calculus in Locally Convex Spaces*. Lecture Notes in Mathematics, vol. 417. Springer, Berlin, New York.

Khesin, B., and Wendt, R. 2009. *The Geometry of Infinite-Dimensional Groups*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. Third Series. A Series of Modern Surveys in Mathematics], vol. 51. Springer, Berlin.

Klingenberg, W. P. A. 1995. *Riemannian Geometry*. Second ed. De Gruyter Studies in Mathematics, vol. 1. Walter de Gruyter & Co., Berlin.

Kolev, B. 2008. Geometric differences between the Burgers and the Camassa–Holm equations. *J. Nonlinear Math. Phys.*, **15**(suppl. 2), 116–132.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Kriegl, A., and Michor, P. W. 1997. *The Convenient Setting of Global Analysis*. Mathematical Surveys and Monographs, vol. 53. American Mathematical Society, Providence, RI.

Lang, S. 1999. *Fundamentals of Differential Geometry*. Graduate Texts in Mathematics, vol. 191. Springer, New York.

Larotonda, G. 2019. Metric geometry of infinite-dimensional Lie groups and their homogeneous spaces. *Forum Math*., **31**(6), 1567–1605.

Le Donne, E., and Züst, R. 2021. Space of signatures as inverse limits of Carnot groups. *ESAIM, Control Optim. Calc. Var*., **27**, 14. Id/No 37.

Lee, J. M. 2013. *Introduction to Smooth Manifolds*. Second ed. Graduate Texts in Mathematics, vol. 218. Springer, New York.

Lenells, J. 2007. The Hunter–Saxton equation describes the geodesic flow on a sphere. *J. Geom. Phys*., **57**(10), 2049–2064.

Lenells, J. 2008. The Hunter–Saxton equation: a geometric approach. *SIAM J. Math. Anal*., **40**(1), 266–277.

Lindenstrauss, J., and Tzafriri, L. 1971. On the complemented subspaces problem. *Isr. J. Math*., **9**, 263–269.

Lyons, T., and Victoir, N. 2007. An extension theorem to rough paths. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, **24**(5), 835–847.

Lyons, T. J. 1998. Differential equations driven by rough signals. *Rev. Mat. Iberoam*., **14**(2), 215–310.

Mackenzie, K. C. H. 2005. *General Theory of Lie Groupoids and Lie Algebroids*. London Mathematical Society Lecture Note Series, vol. 213. Cambridge University Press, Cambridge.

Maeda, Y., Rosenberg, S., and Torres-Ardila, F. 2015. The geometry of loop spaces. I: H^s-Riemannian metrics. *Int. J. Math*., **26**(4), 26. Id/No 1540002.

Magnani, V., and Tiberio, D. 2020. A remark on vanishing geodesic distances in infinite dimensions. *Proc. Am. Math. Soc*., **148**(8), 3653–3656.

Manchon, D. 2008. Hopf algebras in renormalisation. Pages 365–427 of: *Handbook of Algebra. Volume 5*. Elsevier/North-Holland, Amsterdam.

Margalef-Roig, J., and Outerelo Domínguez, E. 1992. *Differential Topology*. North-Holland Publishing Co., Amsterdam.

Marquis, T., and Neeb, K.-H. 2018. Half-Lie groups. *Transform. Groups*, **23**(3), 801–840.

Maurelli, M., Modin, K., and Schmeding, A. 2019. *Incompressible Euler Equations with Stochastic Forcing: A Geometric Approach*. arXiv:1909.09982

McAlpin, J. 1965. *Infinite-Dimensional Manifolds and Morse Theory*. PhD thesis, Columbia University.

Meinrencken, E. 2017. *Lie Groupoids and Lie Algebroids*. www.math.toronto.edu/mein/teaching/MAT1341_LieGroupoids/Groupoids.pdf

Meise, R., and Vogt, D. 1997. *Introduction to Functional Analysis*. Oxford Graduate Texts in Mathematics, vol. 2. The Clarendon Press, Oxford University Press, New York. Translated from the German by M. S. Ramanujan and revised by the authors.

Meyer, R., and Zhu, C. 2015. Groupoids in categories with pretopology. *Theory Appl. Categ*., **30**, 1906–1998.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Micheli, M., Michor, P. W., and Mumford, D. 2013. Sobolev metrics on diffeomorphism groups and the derived geometry of spaces of submanifolds. *Izv. Ross. Akad. Nauk Ser. Mat.*, **77**(3), 109–138.

Michor, P. W., and Vizman, C. 1994. *n*-transitivity of certain diffeomorphism groups. *Acta Math. Univ. Comenian. (N.S.)*, **63**(2), 221–225.

Michor, P. W. 1980. *Manifolds of Differentiable Mappings*. Shiva Mathematics Series, vol. 3. Shiva Publishing, Nantwich.

Michor, P. W. 2020. Manifolds of mappings for continuum mechanics. Pages 3–75 of: *Geometric Continuum Mechanics*, Reuven Segev and Marcelo Epstein (eds). Advances in Continuum Mechanics, vol. 42. Birkhäuser, Boston, MA.

Michor, P. W., and Mumford, D. 2006. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)*, **8**(1), 1–48.

Milnor, J. 1982. *On Infinite-Dimensional Lie Groups*. Preprint, Institute for Advanced Study, Princeton.

Milnor, J. 1984. Remarks on infinite-dimensional Lie groups. Pages 1007–1057 of: *Relativity, Groups and Topology, II (Les Houches, 1983)*. North-Holland, Amsterdam.

Mio, W., Srivastava, A., and Joshi, S. 2007. On shape of plane elastic curves. *Int. J. Comput. Vis.*, **73**, 307–324.

Modin, K. 2019. Geometric hydrodynamics: from Euler, to Poincaré, to Arnold. Pages 71–91 of: *13th Young Researchers Workshop on Geometry, Mechanics and Control. Three Mini-Courses, Coimbra, Portugal, December 6–13, 2018*. Universidade de Coimbra, Departamento de Matemática, Coimbra.

Moerdijk, I., and Pronk, D. A. 1997. Orbifolds, sheaves and groupoids. *K-Theory*, **12**(1), 3–21.

Moerdijk, I., and Mrčun, J. 2003. *Introduction to Foliations and Lie Groupoids*. Cambridge University Press, Cambridge.

Müller, O. 2008. A metric approach to Fréchet geometry. *J. Geom. Phys.*, **58**(11), 1477–1500.

Murua, A., and Sanz-Serna, J. M. 2017. Word series for dynamical systems and their numerical integrators. *Found. Comput. Math.*, **17**(3), 675–712.

Neeb, K.-H. 2005. *Monastir Summer School: Infinite-Dimensional Lie Groups*. Third cycle. Monastir (Tunisie).

Neeb, K.-H. 2006. Towards a Lie theory of locally convex groups. *Jpn. J. Math.*, **1**(2), 291–468.

Neeb, K.-H. 2010. On differentiable vectors for representations of infinite dimensional Lie groups. *J. Funct. Anal.*, **259**(11), 2814–2855.

Omori, H. 1974. *Infinite Dimensional Lie Transformation Groups*. Springer, Cham.

Omori, H. 1978. On Banach–Lie groups acting on finite-dimensional manifolds. *Tôhoku Math. J. II Ser.*, **30**, 223–250.

Omori, H. 1981. A remark on non-enlargeable Lie algebras. *J. Math. Soc. Japan*, **33**(4), 707–710.

Palais, R. S. 1957. *A Global Formulation of the Lie Theory of Transformation Groups*. American Mathematical Society (AMS), Providence, RI.

Palais, R. S. 1968. *Foundations of Global Non-linear Analysis*. Mathematics Lecture Note Series. W. A. Benjamin, New York, Amsterdam.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Poincaré, H. 1901. Sur une forme nouvelle des équations de la mécanique. *C. R. Acad. Sci., Paris*, **132**, 369–371.

Pressley, A., and Segal, G. 1986. *Loop Groups*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, Oxford.

Qian, Z., and Tudor, J. 2011. Differential structure and flow equations on rough path space. *Bull. Sci. Math.*, **135**(6–7), 695–732.

Ree, R. 1958. Lie elements and an algebra associated with shuffles. *Ann. Math. (2)*, **68**, 210–220.

Reutenauer, C. 1993. *Free Lie Algebras*. The Clarendon Press, Oxford University Press, Oxford.

Rudin, W. 1991. *Functional Analysis*. Second ed. McGraw-Hill, Inc., New York.

Rybicki, T. 1995. Isomorphisms between groups of diffeomorphisms. *Proc. Am. Math. Soc.*, **123**(1), 303–310.

Rybicki, T. 2002. A Lie group structure on strict groups. *Publ. Math.*, **61**(3–4), 533–548.

Schmeding, A. 2020. The Lie group of vertical bisections of a regular Lie groupoid. *Forum Math.*, **32**(2), 479–489.

Schmeding, A, and Wockel, C. 2015. The Lie group of bisections of a Lie groupoid. *Ann. Global Anal. Geom.*, **48**(1), 87–123.

Schmeding, A., and Wockel, C. 2016. (Re)constructing Lie groupoids from their bisections and applications to prequantisation. *Differential Geom. Appl.*, **49**, 227–276.

Schmid, R. 2010. Infinite-dimensional Lie groups and algebras in mathematical physics. *Advances in Mathematical Physics*. Vol. 2010, Article ID 280362. https://doi.org/10.1155/2010/280362

Sergeev, A. G. 2020. In search of infinite-dimensional Kähler geometry. *Russ. Math. Surv.*, **75**(2), 321–367.

Sharpe, R. W. 1997. *Differential Geometry: Cartan's Generalization of Klein's Erlangen Program*. Springer, Berlin.

Smolentsev, N. K. 2007. Diffeomorphism groups of compact manifolds. *J. Math. Sci., New York*, **146**(6), 6213–6312.

Stacey, A. 2008. *How to Construct a Dirac Operator in Infinite Dimensions*. arXiv:0809.3104

Steenrod, N. E. 1967. A convenient category of topological spaces. *Michigan Math. J.*, **14**, 133–152.

Takens, F. 1979. Characterization of a differentiable structure by its group of diffeomorphisms. *Bol. Soc. Bras. Mat.*, **10**(1), 17–25.

Tapia, N., and Zambotti, L. 2020. The geometry of the space of branched rough paths. *Proc. Lond. Math. Soc. (3)*, **121**(2), 220–251.

Taylor, M. E. 2011. *Partial Differential Equations I. Basic Theory*. Second ed. Applied Mathematical Sciences, vol. 115. Springer, New York.

Thompson, D. W. 1942. *On Growth and Form*. New edition. Cambridge University Press, Cambridge.

Treves, F. 2006. *Topological Vector Spaces, Distributions and Kernels*. Dover Publications, Inc., Mineola, NY. Unabridged republication of the 1967 original.

Vizman, C. 2008. Geodesic equations on diffeomorphism groups. *SIGMA, Symmetry Integrability Geom. Methods Appl.*, **4**, paper 030, 22.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Voigt, J. 1992. On the convex compactness property for the strong operator topology. *Note Mat*., **12**, 259–269.
Walter, B. 2012. Weighted diffeomorphism groups of Banach spaces and weighted mapping groups. *Dissertationes Math*., **484**, 128.
Werner, D. 2000. *Funktionalanalysis*. Extended ed. Springer, Berlin.
Whitney, H. 1934. Analytic extensions of differentiable functions defined in closed sets. *Trans*. *Amer. Math*. *Soc*., **36**(1), 63–89.
Wockel, C. 2006. *Infinite-Dimensional Lie Theory for Gauge Groups*. PhD thesis, TU Darmstadt.
Wockel, C. 2014. *Infinite-Dimensional and Higher Structures in Differential Geometry*. unpublished lecture notes.
Wockel, C. 2007. Lie group structures on symmetry groups of principal bundles. *J*. *Funct*. *Anal*., **251**(1), 254–288.
Wurzbacher, T. 1995. Symplectic geometry of the loop space of a Riemannian manifold. *J*. *Geom*. *Phys*., **16**(4), 345–384.
Young, L. C. 1936. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math*., **67**, 251–282.

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Index

Banach manifold, 24
Bastiani calculus, 7, 203
Boman's theorem, 202
bornology, 202
bump function, 195
Butcher group, 185
Butcher–Connes–Kreimer algebra, 182

canonical manifold of mappings, 41, 219
carrier of a function, 195
chain rule, 11, 19, 21
change of charts, 16
character group of a Hopf algebra, 178
character of an algebra, 177
Chen's relation, 160, 167, 168
Christoffel symbol, 91
CIA
  see continuous inverse algebra, 49
compact open C^∞-topology, 4, 30, 192, 198, 221
compact open topology, 13, 208
continuous inverse algebra, 49
  unit group, 49
continuous projection, 194
convenient calculus, 202
convenient smooth map, 202
coordinate projection, 11
cotangent bundle, 21, 81
current group, 72
curvature, 96, 100
  Bianchi identity, 101
  sectional, 151
curve
  c-parallel, 99
  energy of a, 84, 222
  length of a, 84
  passing through a point, 19
  piecewise C^1, 84
  principal part of a, 99
  smooth, 2

derivation of an algebra, 227, 230
derivative
  covariant, 94
  covariant, along a map, 109
  directional, 7
  iterated directional, 8
  left logarithmic, 62, 238
  Lie derivative, 227, 236
  metric derivative, 97
  of a manifold-valued curve, 22
diffeomorphism group, 35, 50, 120
  acting n-transitive, 121
  canonical action, ix, 51
  contactomorphisms, 52
  stabiliser of a point, 121
  symplectomorphisms, 52
  volume-preserving, 52, 144, 242
differential form, 231, 240
  (locally) integrable, 237
  Lie derivative, 236
  preserved by a diffeomorphism, 52
  pullback, 235
  volume form, 241
dual space, 4, 83, 193

embedding, 23
  set of, 34
equation of Lie type, 60
Euler equations
  Eulerian form, 143
  Lagrangian form, 145

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Euler–Arnold equation, 152
Euler–Poincaré equation, 149
evaluation map, 39, 210
evolution map, 62
exponential law, 39, 211
exterior differential, 232

fibre product (of manifolds), 27
final topology, 206
flip of the double tangent bundle, 215
formal power series, 183
Fréchet calculus, 14
Fréchet space, 3, 198
Fundamental theorem of calculus, 5

geodesic, 88, 100, 139
  of a spray, 92
  of the L^2-metric, 112
geodesic distance, 85
  non-vanishing, 86
geodesic equation, 92, 99
gradient, 143, 242
Grossman's ellipsoid, 103
group of bisections, 126
  vertical bisections, 130
group of gauge transformations, 77

Hölder continuity, 169
Hahn–Banach theorem, 4
half-Lie group, 155
Hilbert sphere, 16, 83
homogeneous space, 77, 183
Hopf algebra, 177
Hopf–Rinow theorem, 102
Hunter–Saxton equation, 153

immersion, 23
  naïve, 24
  set of, 33
infinitesimally injective (or surjective), 24
initial topology, 206

Kolmogorov's normability criterion, 192

Lie algebra, 55
  associated to a Lie group, 57
  current algebra, 72
  Jacobi identity, 55
  locally convex, 56
  of (left-)invariant vector fields, 56
  of divergence-free vector fields, 243
  of vector fields, 228
Lie bracket, 55
  of vector fields, 56
  trivial, 56
Lie group, 48
  Bourbaki construction principle, 70
  exponential, 65
  exponential coordinates, 66
  left- (right-)translation, 48
  Lie subgroup, 52
  locally exponential, 66
  regular (in the sense of Milnor), 62
  semidirect product, 54
  semiregular, 62
  split exact sequence, 54
  tangent Lie group, 53
Lie group action, 51
  adjoint action, 60
Lie groupoid, 125
  action groupoid, 127
  bisection action groupoid, 134
  current groupoid, 128
  enough bisections, 135
  gauge groupoid, 128
  pair groupoid, 127
  tangent groupoid, 132
  unit groupoid, 127
  vertex group, 125
Lie polynomial, 164
Lie series, 164
Lie theorems, 60
local addition, 43, 213
  normalised, 218
local-to-global argument, 196
loop group, 74
Lyons lift, 170
Lyons' lifting theorem, 170

Mackey complete, 5, 187, 202
manifold atlas, 16
manifold chart, 16
Maurer–Cartan form, 237
mean value theorem, 9
Minkowski functional, 83, 192
multiplicative functional, 168, 179

Neumann inversion formula, 161

Omori's theorem, ix, 155
orbifold, 124, 127

partition of unity, 195
pre-shape space, 114

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

principal bundle, 76, 129
  gauge group, 76
  structure group, 76
pro-Lie group, 166
product manifold, 17
pullback, 31, 45, 209
pushforward, 13, 31, 43, 209

Riemann–Stieltjes sum, 157, 158
Riemannian metric
  H^1-metric, 113
  $\dot{H}^1$-semimetric, 153
  L^2-metric, 81, 108, 140
  co-orthogonal structure, 102
  elastic metric, 116
  invariant L^2-metric, 83, 88, 89, 144
  pullback metric, 116
  right/left invariant, 87
  robust, 101
  strong, 81
  weak, 80
rough norm, 197
rough path, 170
  branched, 181, 183
  Brownian motion, 171
  set of weakly geometric, 170
  weakly geometric, 170
rule on partial differentials, 9, 22

Schwartz' theorem, 8, 12
seminorm, 3
  basis condition, 191
  fundamental system of, 191
  generating family of, 3
  separating family of, 190
shape analysis, 106
shape space, 114
shuffle algebra, 176, 183
signature of a smooth path, 162
smooth variation, 138
  right-shifted, 146
space
  C^k-paracompact, 196
  C^k-regular, 196
  complemented subspace, 194
  convenient vector space, 202
  locally convex, 3, 192
  metrisable, 3
  of Hölder continuous functions, 169
  of smooth p-forms, 232
  sequentially closed, 12
  short exact sequence, 25
  topological vector space, 1, 186
spray, 91
  associated covariant derivative, 95
  connector, 96
  metric, 92
  metric spray, 97
square root velocity transform, 116
SRVT
  see square root velocity transform, 116
Stacey–Roberts Lemma, 45
standard simplex (Δ), 168
submanifold, 17
  split, 17
submersion, 23
  naïve, 24
  set of, 33
support of a function, 195

tangent bundle, 19
tangent map, 21
tangent space, 19
tangent vector (geometric), 19
tensor algebra, 160
topology of compact convergence, 208
transversal, 26

unit group (of an algebra), 49

vector bundle, 215
  direct product, 217
  dual bundle, 21, 81
  pullback, 216
  set of all sections, 215
  smooth section, 215
  Whitney sum, 217, 220
vector bundle morphism, 215
vector field, 225
  (left) invariant, 56, 64
  along a smooth map, 109
  complete, 64
  divergence, 143, 242
  flow, 226
  Helmholtz decomposition, 242
  integral curve, 226
  Lie bracket, 226
  local representative, 225
  principal part, 225

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

related, 226
second-order, 91
vector space of, 225
vector topology, 1
generated by seminorms, 3
Volterra series, 63, 162

weak integral, 4
wedge product, 234

Young integral, 158

zero-section, 43, 213, 220

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC

Downloaded from https://www.cambridge.org/core. IP address: 189.217.92.64, on 12 Feb 2026 at 23:12:31, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://www.cambridge.org/core/product/6483795C98EE417C0F3654F6C192C3BC